Second-order asymptotics for the block counting process in a class of regularly varying Λ-coalescents

VLADA LIMIC*
CNRS UMR 8628, Laboratoire
de Mathématiques, Université
Paris-Sud, Bâtiment 425
91405 Orsay, France

e-mail: vlada@cmi.univ-mrs.fr

ANNA TALARCZYK†
Institute of Mathematics
University of Warsaw
ul. Banacha 2, 02-097 Warszawa
Poland

e-mail: annatal@mimuw.edu.pl

Abstract

Consider a standard Λ-coalescent that comes down from infinity. Such a coalescent starts from a configuration consisting of infinitely many blocks at time 0, but its number of blocks N_t is a finite random variable at each positive time t. Berestycki et al. (2010) found the first-order approximation v for the process N at small times. This is a deterministic function satisfying $N_t/v_t \to 1$ as $t \to 0$. The present paper reports on the first progress in the study of the second-order asymptotics for N at small times. We show, that if the driving measure Λ has a density near zero, which behaves as $x^{-\beta}$ with $\beta \in (0,1)$ then the process $(\varepsilon^{-1/(1+\beta)}(N_{\varepsilon t}/v_{\varepsilon t}-1))_{t\geq 0}$ converges in law as $\varepsilon \to 0$ in the Skorokhod space to a totally skewed $(1+\beta)$-stable process. Moreover, this process is a unique solution of a related stochastic differential equation of Ornstein-Uhlenbeck type, with a completely asymmetric stable Lévy noise.

Keywords: Λ-coalescent, coming down from infinity, second-order approximations, stable Lévy process, Ornstein-Uhlenbeck process, Poisson random measure

AMS 2000 subject classifications: Primary 60J25; Secondary 60F17, 92D25, 60G52, 60G55

*Research supported in part by the ANR MANEGE grant.
†Research supported in part by MNiSzW grant N N201 397537 (Poland).

1 Introduction

The Λ-coalescents were introduced and first studied independently by Pitman [19] and Sagitov [20] and are also considered in a contemporaneous work of Donnelly and Kurtz [12]. They are useful models of genealogical trees of populations that evolve under the assumption of unbounded variance in the reproduction (resampling) mechanism. Berestycki et al. [3] derive the first order approximation for the number of blocks in a general standard Λ-coalescent that comes down from infinity. The present work initiates the study of the second order approximation for the same process. We next recall the basic definitions, mention some of the landmark results, and present the motivation for the problem we resolved in this work. For recent overviews of the literature we refer the reader to [6, 2].

Let Λ be an arbitrary finite measure on $[0, 1]$. We denote by $(\Pi_t, t \geq 0)$ the associated Λ-coalescent. This Markov jump process $(\Pi_t, t \geq 0)$ takes values in the set of partitions of $\{1, 2, \ldots\}$. Its law is specified by the requirement that, for any $n \in \mathbb{N}$, the restriction Π^n of Π to $\{1, \ldots, n\}$ is a continuous-time Markov chain with the following transitions: whenever Π^n has $b \in [2, n]$ blocks, any given k-tuple of blocks coalesces at rate $\lambda_{b,k} := \int_{[0,1]} r^{k-2}(1-r)^{b-k}\Lambda(dr)$. The total mass of Λ can be scaled to 1. This is convenient for the analysis, and corresponds to a constant time rescaling of the process. Henceforth we assume that Λ is a probability measure.

The *standard* Λ-coalescent starts from the trivial configuration $\{\{i\} : i \in \mathbb{N}\}$. Let us denote by $N^\Lambda(t)$ (or $N(t)$ if clear from the context) the number of blocks of $\Pi(t)$ at time t. If $\mathbb{P}(N^\Lambda(t) < \infty, \forall t > 0) = 1$ the coalescent is said to *come down from infinity.* As part of his thesis work, Schweinsberg [22] derived the following criterion: the (standard) Λ-coalescent *comes down from infinity* (CDI) if and only if

$$\sum_{b=2}^{\infty} \left(\sum_{k=2}^{b} (k-1) \binom{b}{k} \lambda_{b,k} \right)^{-1} < \infty. \tag{1.1}$$

Let

$$\Psi^*(q) = \int_0^1 (qy - 1 + e^{-yq})) \frac{\Lambda(dy)}{y^2}. \tag{1.2}$$

Bertoin and Le Gall [7] obtained an equivalent condition: Λ-coalescent CDI if and only if

$$\int_a^\infty \frac{1}{\Psi^*(q)} dq < \infty, \text{ for some (and then all) } a > 0. \tag{1.3}$$

Throughout the paper we will assume (1.3). Let $N = (N_t, t \geq 0)$ be the block counting process defined above, so that $N(0) = \infty$ and $\mathbb{P}(N_t < \infty, t > 0) = 1$. As indicated above, in [3], Theorem 1 it is shown that, solely under (1.3), there exists a "law of large numbers" approximation for the block counting process, more precisely,

$$\lim_{t \to 0+} N_t / v_t^* = 1, \text{ almost surely,} \tag{1.4}$$

where v^* is uniquely determined by $\int_{v_t^*}^\infty \frac{1}{\Psi^*(q)}dq = t$, for all $t > 0$. Any function satisfying (1.4) is called a *speed of coming down from infinity*, or a *speed of CDI*.

Instead of Ψ^* we choose to work with $\Psi : [1,\infty) \mapsto \mathbb{R}_+$ defined by

$$\Psi(q) = \int_0^1 (qy - 1 + (1-y)^q)\frac{\Lambda(dy)}{y^2}. \tag{1.5}$$

This function is different from Ψ used in [3] (which is now our Ψ^*). Moreover, our Ψ appeared as $\bar{\Psi}$ in [5, 15, 16, 17] where it was already noted that this function arises from the model in a more natural way, and it may be more convenient for analysis than Ψ^*. It is not difficult to see that Ψ and Ψ^* have the same asymptotic behavior at ∞ (see Lemma 2.1 or [16, 5]), and that therefore (1.1)–(1.3) are further equivalent to

$$\int_a^\infty \frac{1}{\Psi(q)}dq < \infty, \text{ for some (and then all) } a > 1. \tag{1.6}$$

Moreover, if we define $v : \mathbb{R}_+ \mapsto \mathbb{R}_+$ by

$$t = \int_{v_t}^\infty \frac{1}{\Psi(q)}dq. \tag{1.7}$$

then (see Lemma 2.2) $v_t \sim v_t^*$ as $t \to 0$, and so v is also a speed of CDI for the corresponding Λ-coalescent.

From the results of Berestycki et al. [3] it follows that the asymptotic behavior of the speed of CDI v_t for small t depends very strongly on the behavior of the driving measure Λ near 0. This is caused by the fact that the behavior of Λ near 0 is linked to the asymptotics of $\Psi(q)$ as $q \to \infty$ by a result of a tauberian nature. For example, if for small x

$$\Lambda(dx) \approx x^{-\beta}dx, \qquad \text{with } \beta \in (0,1), \tag{1.8}$$

then $v_t \sim C\, t^{-\frac{1}{\beta}}$, for some $C \in (0,\infty)$, as $t \to 0$ (see Lemma 5.1). Note that (1.8) is understood in the sense of Assumption (A) in Section 4.

A natural question is to study the second-order fluctuations of N about its speed of CDI. In particular, one wishes to understand how close is $\frac{N_t}{v_t}$ to 1 at small times, and if this proximity can be measured in some regular (and universal) way. In the present paper we address this problem by considering the fluctuations in a functional sense, with time scaled by $\varepsilon \to 0$. More precisely, we investigate the convergence in law of the processes

$$\left(r(\varepsilon)\left(\frac{N_{\varepsilon t}}{v_{\varepsilon t}} - 1\right), t \geq 0 \right), \tag{1.9}$$

were $r(\varepsilon)$ is an appropriately chosen normalization so that the limit process is non trivial.

It turns out that both the normalization $r(\varepsilon)$ and the limit process again depend on the behavior of Λ near 0. The singularity exponent β of the density of Λ near 0 decides the rate of convergence of $\frac{N_t}{v_t^*}$, and therefore of $\frac{N_t}{v_t}$, to 1.

Our main result (see Theorem 4.2) states that if Λ satisfies (1.8), then the normalization yielding a non trivial limit of (1.9) is $r(\varepsilon) = \varepsilon^{-\frac{1}{1+\beta}}$, and the processes defined by (1.9) converge in law in the Skorokhod space $D([0,\infty))$ (with the usual J_1 topology) to a $(1+\beta)$-stable process which is totally skewed to the left. This process can be further described as a unique solution of a stochastic differential equation of an Ornstein-Uhlenbeck type with Lévy noise (see (4.5)).

Theorem 4.2 can be applied to the important class of Beta-coalescents. Moreover, condition (1.8) implies that our Λ-coalescent is in a sense "close" to a Beta coalescent. The fact that the limit processes in (1.9) are $(1+\beta)$-stable processes can be explained by the fact that Beta-coalescents that come down from infinity, i.e. satisfying (1.8) with $0 < \beta < 1$ are related to genealogies of populations with infinite variance branching (see Sagitov [20] and Schweinsberg [24]). On the other hand, limits of fluctuations related to such infinite variance branching systems are usually stable processes (with index of stability depending on the type of the branching, see e.g. [14, 10]). Furthermore, Birkner et al. [9] show relations between continuous state $(1+\beta)$-stable branching processes and Beta coalescents satisfying (1.8). See also Remark 4.3(c).

In Theorem 4.4(a) we prove that one can replace v by v^* in (1.9) and obtain exactly the same result as for v.

Under assumption (1.8) both v_t and $v_t^* \sim w_t = Ct^{-\frac{1}{\beta}}$. It is therefore natural to ask whether one obtains the same results if in (1.9) v is replaced by w. The answer is positive only if one assumes additional regularity of Λ near 0 (see Theorem 4.4(b)). Furthermore, we prove that if the density of Λ does not satisfy this additional regularity condition, then the natural candidate for speed w may not behave well with respect to the path approximation (see Remark 4.5 and Section 8.2.) Due to this lack of universality/robustness, one can conclude that the second-order approximation to N is a more subtle problem than that of finding the speed of CDI.

It is interesting to note that the present analysis (in the sense of functional convergence) was not carried out even for the case of the Kingman coalescent, where Λ is the Dirac measure at 0. It is known, that in this case the law of $t^{-\frac{1}{2}}(N_t/v_t - 1)$ converges to a Gaussian law, see for example Aldous [1]. Here we assume that $\Lambda(\{0\}) = 0$, so that the coalescent does not have the Kingman part. We postpone the study of the complementary setting to a future work. We conjecture that in the case of the pure Kingman coalescent (i.e. Λ is the Dirac mass at 0) the limit process in (1.9) will have a form similar to (4.3), where the integration with respect to the stable random measure is replaced by integration with respect to a Brownian motion. The Kingman case, although seemingly easier can not be done with our present technique, since here we rely heavily on the Poisson process construction of a Λ coalescent, which is particularly nice if $\Lambda(\{0\}) = 0$. We also make the usual assumption $\Lambda(\{1\}) = 0$, under which the Λ-coalescent either comes down from infinity or stays infinite forever (see Pitman [19]).

When $\Lambda(\{0\}) = 0$, one can construct a realization of the corresponding Λ-coalescent from a Poisson point process in the following (now standard)

way. Let

$$\pi(\cdot) = \sum_{i \in \mathbb{N}} \delta_{(T_i, Y_i)}(\cdot) \tag{1.10}$$

be a Poisson point process on $\mathbb{R}_+ \times (0,1)$ with intensity measure $dt \otimes \nu(dy)$ where $\nu(dy) = y^{-2}\Lambda(dy)$. Each atom (t, y) of π impacts the evolution of Π as follows: for each block of $\Pi(t-)$ a coin is flipped with probability of heads equal to y; all the blocks corresponding to coins that come up "head" are merged immediately into one single block, and all the other blocks remain unchanged. In order to make this construction rigorous, one initially considers the restrictions $(\Pi^{(n)}(t), t \geq 0)$, since the measure ν may be infinite (see for example [2, 6]).

Our technique is based on a novel approach using an explicit representation of the block counting process in terms of an enriched Poisson random measure π^E, which is defined on a larger space in such a way that it also includes the information on (individual block) coloring. One can then write an integral equation for the number of blocks N_t involving an integral with respect to π^E. This equation turns out to be analytically tractable. In our approach, we rely on the properties of integrals with respect to Poisson, compensated Poisson and stable random measures, Laplace transforms of Poisson integrals and of totally skewed stable random variables, as well as standard tools in the analysis of processes in the Skorokhod space, e.g. the Aldous criterion for tightness. Moreover, a deterministic lemma from [3], for comparing solutions to two different Cauchy (or Cauchy-like) problems, turns out to be very useful.

The remainder of the paper is organized as follows. In Section 2 we give some basic information on general Λ coalescents; in Section 3 (still in the general setting) we develop the integral equations for N and N/v and study their basic properties; Section 4 contains the main results of the paper; Section 5 contains proofs of the results from the Section 3, as well as an additional Lemma on the properties of Ψ and v under assumption (1.8); Section 6 contains the proofs remaining from Section 3; in Section 7 we give the proof of the main result – Theorem 4.2; in Section 8 we prove Theorem 4.4 and discuss the problem of robustness.

Throughout the paper $C, C_1, C_2, \ldots$ always denote positive constants which may be different from line to line.

2 Background and preliminary results

In this section we collect some of the basic properties of Ψ and v and their relation to the block counting process N. The proofs are given in section 5.

Recall that Ψ and v are defined by (1.5) and (1.7), respectively. Let us also define

$$h(q) := \frac{\Psi(q)}{q}. \tag{2.1}$$

For $0 < a \leq 1$ let Ψ_a (resp. Ψ_a^*) be defined by (1.5) (resp. (1.2)) with $\Lambda(dy)$ replaced by $\mathbb{1}_{[0,a]}(y)\Lambda(dy)$.

The first lemma concerns the most general setting, up to time-change.

Lemma 2.1. *Let Λ be an arbitrary probability measure on $[0,1]$ satisfying $\Lambda(\{0\}) = \Lambda(\{1\}) = 0$. Then the function Ψ given by (1.5) is well defined on $[1,\infty)$. In addition,*
(i) Ψ is continuous on $[1,\infty)$ and strictly positive on $(1,\infty)$,
(ii) for any $q \geq 1$

$$\Psi(q) \leq q(q-1), \tag{2.2}$$

$$0 \leq \Psi^*(q) - \Psi(q) \leq \frac{q}{2}, \tag{2.3}$$

(iii) for any $q \geq 1$ and $a \in (0,1)$

$$0 \leq \Psi(q) - \Psi_a(q) \leq \frac{q}{a^2}, \tag{2.4}$$

$$0 \leq \Psi^*(q) - \Psi_a^*(q) \leq \frac{q}{a^2}, \tag{2.5}$$

(iv) and both Ψ and h are strictly increasing on $[1,\infty)$ and differentiable on $(1,\infty)$.

Most of these facts are known in the literature but for the benefit of the reader we will include a short proof in Section 5. Note that (2.3) implies the equivalence of (1.3) and (1.6).

From now on assume that $\Lambda(\{0\}) = \Lambda(\{1\}) = 0$ and that the Λ-coalescent comes down from infinity, which is equivalent to any (and therefore all) of (1.1), (1.3), (1.6). By Lemma 2.1, Ψ is a continuous and strictly increasing function on $[1,\infty)$, strictly positive on $(1,\infty)$, and $\int_1^\infty dq/\Psi(q) \geq \int_1^\infty dq/q(q-1) = \infty$. This implies that v is a well defined strictly decreasing function on $(0,\infty)$. Moreover, v has the following properties.

Lemma 2.2. *(i) $v_t > 1$ for all $t > 0$, $\lim_{t\to 0+} v_t = \infty$ and $\lim_{t\to\infty} v_t = 1$,*
(ii) v is differentiable and

$$v_t' = -\Psi(v_t), \tag{2.6}$$

(iii) and

$$\lim_{t\to 0+} \frac{v_t}{v_t^*} = 1. \tag{2.7}$$

(iv) Therefore

$$\lim_{t\to 0+} \frac{N_t}{v_t} = 1 \quad \textit{almost surely}, \tag{2.8}$$

(v) and for any $p > 0$,

$$\lim_{t\to 0+} E \sup_{0<s\leq t} \left| \frac{N_s}{v_s} - 1 \right|^p = 0. \tag{2.9}$$

Moreover, for any $p > 0$ there exists $C_p > 0$ such that

$$E \sup_{s\geq 0} \left(\frac{N_s}{v_s} \right)^p \leq C_p. \tag{2.10}$$

Remark 2.3. Lemma 2.2 parts (iv) and (v) say that $\frac{N_t}{v_t}$ converges to 1 almost surely and in L^p, for any $p > 0$. This was shown with v^* in place of v in [3] Theorems 1 and 2. Moreover, in the same article a (2.10) was derived,

again with v^* in place of v. (Note that [3] Theorem 2 assumes that $p \geq 1$, but this can be easily extended to all $p \in (0,1)$ by Jensen's inequality.) Due to (2.7), we can thus obtain (iv)–(v) without any additional work. In return, Lemma 3.6 stated at the end of Section 3 is a novel and stronger estimate, important for our analysis. $\square$

Let us also recall the following elementary estimate, that will be used frequently in the proofs (see [3], Lemma 10 for derivation).

Lemma 2.4. *Suppose $f, g : [a,b] \mapsto \mathbb{R}$ are càdlàg functions such that*

$$\sup_{x\in[a,b]} \left| f(x) + \int_a^x g(u)\, du \right| \leq c\,, \tag{2.11}$$

for some $c < \infty$. If in addition $f(x)g(x) > 0$, $x \in [a,b]$ whenever $f(x) \neq 0$, then

$$\sup_{x\in[a,b]} \left| \int_a^x g(u)\, du \right| \leq c \quad \textit{and} \quad \sup_{x\in[a,b]} |f(x)| \leq 2c.$$

3 Integral equations for N

In this section we give a representation of the block counting process N of a Λ-coalescent in terms of an integral equation involving Poisson random measure. We also write an equation for the process N divided by the speed of CDI. Some preliminary estimates are also given. The proofs are postponed until Section 6.

This construction is our starting point to studying fluctuations of the block counting process about the speed of CDI, carried out in the forthcoming sections. The approach presented here is quite general, and we hope it to be of independent interest.

In this section and the rest of the paper we always assume that $\Lambda(\{0\}) = \Lambda(\{1\}) = 0$ and that any (and therefore all) of (1.1), (1.3), (1.6) hold.

As discussed in the Introduction, Λ-coalescent can be constructed via a coloring procedure which is based on a Poisson random measure π on $[0,\infty) \times [0,1]$, and an independent assignment of colors to the blocks. Here we introduce an enriched Poisson random measure which contains all the information on the coloring. This is a key ingredient in the first important novelty of our approach - an explicit representation of the martingale which drives the block counting process N.

In order to explain this now, we will need some additional notation. As usual, let $\mathbb{N}$ denote the set of natural numbers (without zero). Let μ be the law of a sequence of i.i.d. random variables $X_1, X_2, \ldots$ uniformly distributed on $[0,1]$, i.e. μ is a probability measure on $[0,1]^{\mathbb{N}}$, equipped with the product σ-algebra generated by the cylinder sets of the form $B_1 \times B_2 \times \ldots B_n \times [0,1] \times [0,1] \times \ldots$, $n \in \mathbb{N}$, $B_i \in \mathscr{B}([0,1])$, $i \in \mathbb{N}$. The vectors in $[0,1]^{\mathbb{N}}$ will be denoted in boldface $\mathbf{x} = (x_1, x_2, \ldots) \in [0,1]^{\mathbb{N}}$. We will usually write $d\mathbf{x}$ instead of $\mu(d\mathbf{x})$.

Let π^E be a Poisson random measure on $[0,\infty) \times [0,1] \times [0,1]^{\mathbb{N}}$ with intensity measure $ds\frac{\Lambda(dy)}{y^2}d\mathbf{x}$. Observe that such a random measure can be

constructed using a Poisson random measure π from (1.10) and an independent array of i.i.d. random variables $(X_j^i)_{i,j\in\mathbb{N}}$, where X_j^i have uniform distribution on $[0,1]$. Then $\pi^E = \sum_{i\in\mathbb{N}} \delta_{(T_i,Y_i,\mathbf{X^i})}$ is a Poisson random measure with intensity $ds\frac{\Lambda(dy)}{y^2}d\mathbf{x}$.

Moreover, π and π^E are coupled by the relation

$$\pi(\cdot) = \pi^E(\cdot \times [0,1]^{\mathbb{N}}). \tag{3.1}$$

We will henceforth assume that (3.1) holds. Then we can construct the Λ coalescent by the following procedure: upon arrival of an atom $(t, y, \mathbf{x})$ of π^E, the jth block present in the configuration at time $t-$ is colored if and only if $x_j < y$. Once the colors are assigned, to form the configuration at time t, merge all the colored blocks into a single block, and leave the other (uncolored) blocks intact.

Recall that we assume that the coalescent comes down from infinity, so $N_r < \infty$ a.s. for any $r > 0$. The procedure described above implies that

$$N_t = N_r - \int_{(r,t]\times[0,1]\times[0,1]^{\mathbb{N}}} f(N_{s-}, y, \mathbf{x})\pi^E(dsdyd\mathbf{x}), \text{ for all } 0 < r < t, \tag{3.2}$$

where f is a function which quantifies the decrease in the number of blocks during a coalescing event:

$$f(k, y, \mathbf{x}) = \left(\sum_{j=1}^{k} \mathbf{1}_{\{x_i<y\}} - 1\right) \vee 0 = \sum_{j=1}^{k} \mathbf{1}_{\{x_i<y\}} - 1 + \mathbf{1}_{\bigcap_{j=1}^{k}\{x_j\geq y\}}. \tag{3.3}$$

Integration with respect to Poisson random measures is well understood; the reader is referred, for example, to [18].

Recall (1.5). One can easily see that

$$\Psi(k) = \int_{[0,1]\times[0,1]^{\mathbb{N}}} f(k, y, \mathbf{x})\frac{\Lambda(dy)}{y^2}d\mathbf{x}. \tag{3.4}$$

Since Ψ is an increasing function and N a decreasing process, we have

$$\int_{(r,t]} \Psi(N_{s-})ds \leq \Psi(N_r)(t-r) \leq N_r^2(t-r),$$

where the last inequality is due to (2.2). We know that $EN_r^2 < \infty$, (see e.g. (2.10)) hence

$$E\int_{(r,t]\times[0,1]\times[0,1]^{\mathbb{N}}} f(N_{s-}, y, \mathbf{x})ds\frac{\Lambda(dy)}{y^2}d\mathbf{x} < \infty.$$

This implies that the integral in (3.2) belongs to L^1 (see e.g. Theorem 8.23 in [18]).

As the first step towards the proof of Theorem 4.2 we have just shown (see (3.2) and (3.4)) that

Lemma 3.1. *For any $0 < r < t$*

$$N_t = N_r - \int_r^t \Psi(N_s)ds - \int_{(r,t]\times[0,1]\times[0,1]^{\mathbb{N}}} f(N_{s-}, y, \mathbf{x})\hat{\pi}^E(dsdyd\mathbf{x}), \quad (3.5)$$

where $\hat{\pi}^E$ denotes the compensated Poisson random measure

$$\hat{\pi}^E(dsdyd\mathbf{x}) = \pi^E(dsdydx) - ds\frac{\Lambda(dy)}{y^2}d\mathbf{x}. \quad (3.6)$$

Remark 3.2. The above representation can be done for $N^{(n)}$, the counting process of the number of blocks of a Λ-coalescent starting from n blocks, even if the Λ-coalescent does not come down from infinity. Moreover, a similar representation exists for Ξ-coalescents, and might be useful in similar type of analysis as done here. For background on this general class of exchangeable coalescents we refer the reader to [23, 2, 6]. □

More importantly, we can write a stochastic integral equation for $\frac{N_t}{v_t}$. Indeed, due to (1.7) we have

$$v_t = v_r - \int_r^t \Psi(v_s)ds, \; 0 < r < t,$$

thus

$$\frac{1}{v_t} = \frac{1}{v_r} + \int_r^t \frac{\Psi(v_s)}{v_s^2}ds$$

and therefore (3.2) and a simple application of the integration by parts yield

Lemma 3.3. *For any $0 < r < t$*

$$\begin{aligned}\frac{N_t}{v_t} = \frac{N_r}{v_r} - \int_r^t \frac{N_s}{v_s}\left(\frac{\Psi(N_s)}{N_s} - \frac{\Psi(v_s)}{v_s}\right)ds \\ - \int_{(r,t]\times[0,1]\times[0,1]^{\mathbb{N}}} \frac{f(N_{s-}, y, \mathbf{x})}{v_s}\hat{\pi}^E(dsdyd\mathbf{x}), \quad (3.7)\end{aligned}$$

where $\hat{\pi}^E$ is as in (3.6).

Remark 3.4. A predecessor of this result existed in [3, 15], where the process of main interest was $\log N/v$ instead of N/v, but the martingale part was never explicitly written down, and therefore it could not be used in such as precise way as it is about to be used here. □

It is natural to continue by investigating the integral with respect to $\hat{\pi}^E$.

Lemma 3.5. *The process $\tilde{M} = (\tilde{M}(t))_{t\geq 0}$, where*

$$\tilde{M}(t) = \int_{[0,t]\times[0,1]\times[0,1]^{\mathbb{N}}} \frac{f(N_{s-}, y, \mathbf{x})}{v_s}\hat{\pi}^E(dsdyd\mathbf{x}) \quad (3.8)$$

is a well defined, square integrable martingale with quadratic variation

$$[\tilde{M}](t) = \int_{[0,t]\times[0,1]\times[0,1]^{\mathbb{N}}} \left(\frac{f(N_{s-}, y, \mathbf{x})}{v_s}\right)^2 \pi^E(dsdyd\mathbf{x}). \quad (3.9)$$

Moreover, for any $p \in (0,2]$ there exists $C_p > 0$, such that for all $t > 0$

$$E \sup_{0 \le s \le t} \left| \tilde{M}(s) \right|^p \le C_p t^{\frac{p}{2}}. \tag{3.10}$$

Using (3.7) and Lemma 3.5 one can improve on (2.9) as follows.

Lemma 3.6. *If the Λ-coalescent comes down from infinity then for any $p \in (0,2]$ there exists $0 < C_p < \infty$ such that*

$$E \sup_{s \le t} \left| \frac{N_s}{v_s} - 1 \right|^p \le C_p t^{p/2}. \tag{3.11}$$

4 Statement of the main result

In this section we formulate the main result, together with several lemmas needed to state the result. All the proofs are postponed until Section 7. We use (often without mention) the notation set in the Introduction.

As explained in the Introduction, we are interested in the rate of convergence in (2.8). It turns out that this rate, as well as the type of the limit process, depends on very fine properties of the driving measure Λ. This may seem surprising, in view of the bound (3.10), which becomes an equality asymptotically as $t \to 0$. In fact (3.10) was already implicit in [3], at least for $p = 2$, where the infinitesimal variance of an analogous martingale (the one driving the equation for $\log \frac{N.}{v.}$) was carefully estimated, even though that martingale was not as explicitly expressed there as $\tilde{M}$ is expressed here. Without paying consideration to the size of jumps of N at small times, these inequalities (which are asymptotically true as equalities) may suggests Gaussian type limits for appropriately rescaled $\tilde{M}$ (and therefore $N/v - 1$). This indeed turns out to be the case in the setting of the Kingman coalescent (not treated here, as remarked in Introduction the reader can check [1] for the non-functional CLT in this setting). However, one quickly realizes that under Assumption (A) the largest jumps of $\tilde{M}$ in $[0,t]$ are not $o(\sqrt{t})$ as $t \to 0$ (in fact they are of order $t^{1/(1+\beta)} >> \sqrt{t}$). Hence the Gaussian scaling is not appropriate and moreover, the limiting process (if there is one) is expected to have jumps. We refer the reader to Remark 4.3(c) for further intuition regarding the limiting process.

Our main result will be derived under the following assumptions.

Assumption. $\Lambda(\{0\}) = \Lambda(\{1\}) = 0$. Moreover, there exists $y_0 \le 1$ such that

$$\Lambda(dy) = g(y)dy, \ y \in [0, y_0] \ \text{ and } \ \lim_{y \to 0+} g(y) y^\beta = A, \tag{A}$$

for some $0 < \beta < 1$ and $0 < A < \infty$.

Remark 4.1. (a) Condition $\beta > 0$ ensures that the Λ coalescent satisfies (1.6), hence that it comes down from infinity, since it is not difficult to see that (A) implies that $\Psi(q) \sim C q^{1+\beta}$ (see also Lemma 5.1 below). Condition $\beta < 1$ is clear, since Λ has to be a finite measure.
(b) Assumption (A) is satisfied by all the Beta-coalescents that come down

from infinity, i.e. all the coalescents where Λ has density of the form $g(y) = \frac{1}{B(1-\beta,\alpha)} y^{-\beta}(1-y)^{\alpha-1}$, for some $0 < \beta < 1$ and $\alpha > 0$ and the normalizing constant is the appropriately evaluated Beta function.
(c) Under Assumption (A) we can obtain precise asymptotics of the speed of coming down from infinity v and the functions Ψ and h, see Lemma 5.1. In particular, as $t \to 0$ we have $v_t \sim v_t^* \sim K_1 t^{-\frac{1}{\beta}}$, where

$$K_1 = \left(\frac{1+\beta}{A\Gamma(1-\beta)} \right)^{\frac{1}{\beta}}, \tag{4.1}$$

and where Γ is the Gamma function. $\square$

We shall study the asymptotic behavior, as $\varepsilon \to 0$, of the process $X_\varepsilon = (X_\varepsilon(t))_{t \geq 0}$ defined by

$$X_\varepsilon(0) = 0 \text{ and } X_\varepsilon(t) = \varepsilon^{-\frac{1}{1+\beta}} \left(\frac{N_{\varepsilon t}}{v_{\varepsilon t}} - 1 \right), \ t > 0. \tag{4.2}$$

For each $B \in \mathscr{B}(\mathbb{R})$ Borel set, let $|B|$ denote its Lebesgue measure. Let $\mathcal{M}$ be an independently scattered $(1+\beta)$-stable random measure on $\mathbb{R}$ with skewness intensity 1. That is, for each $B \in \mathscr{B}(\mathbb{R})$ such that $0 < |B| < \infty$, $\mathcal{M}(B)$ is a $(1+\beta)$-stable random variable with characteristic function

$$\exp\left\{ -|B|\,|z|^{1+\beta} \left(1 - i(\operatorname{sgn} z) \tan \frac{\pi(1+\beta)}{2} \right) \right\}, \ z \in \mathbb{R},$$

$\mathcal{M}(B_1), \mathcal{M}(B_2), \ldots$ are independent whenever $B_1, B_2, \ldots$ are disjoint sets, and $\mathcal{M}$ is σ-additive a.s. (see Samorodnitsky and Taqqu [21], Def. 3.3.1).

We are now ready to state the main result.

Theorem 4.2. *Assuming* (A), *the process* X_ε *defined in* (4.2) *converges in law in the Skorokhod space* $D([0,\infty))$ *to a* $(1+\beta)$*-stable process* $Z = (Z_t)_{t \geq 0}$ *given by*

$$Z(t) = -\frac{K}{t} \int_0^t u \mathcal{M}(du), \ t > 0, \ Z(0) = 0, \tag{4.3}$$

where K *is the following positive constant*

$$K = \left(-A \int_0^\infty (e^{-y} - 1 + y) y^{-2-\beta} dy \cos \frac{\pi(1+\beta)}{2} \right)^{\frac{1}{1+\beta}}. \tag{4.4}$$

Remark 4.3. (a) The integral in (4.3) is understood in the sense of Chapter 3 of [21].
(b) The process Z can be also expressed as

$$Z(t) = -\frac{K}{t} \int_0^t u dL_u,$$

where L is the $(1+\beta)$-stable totally skewed to the right (having no negative jumps) Lévy process. Moreover, Z solves the following stochastic differential equation of the Ornstein-Uhlenbeck type

$$Z(t) = -\int_0^t s^{-1} Z(s)\, ds - KL(t). \tag{4.5}$$

(c) Observe that Assumption (A) is satisfied by Beta-coalescents which come down from infinity. Furthermore, if Λ satisfies Assumption (A), then, from the point of view of behavior of N_t, v_t and $N_t/v_t - 1$ near 0 the Λ-coalescent resembles a corresponding Beta-coalescent (or rather a class of Beta-coalescents) having driving measure(s) of the form Beta$(1-\beta, a)$, for some $a > 0$.

The fact that the limit process is $(1+\beta)$-stable can be explained by observing that for each $\beta \in (0,1)$, one member of the above family (notably the Beta$(1-\beta,\beta)$-coalescent) was obtained from genealogies of populations with supercritical infinite variance branching both by Sagitov [20] (in his setting, the branching mechanism has generating function $1 - \frac{1+\beta}{\beta}(1-s) + \frac{1}{\beta}(1-s)^{1+\beta}$) and by Schweinsberg [24] (in his setting, the probability that the individual has k or more offspring decays like $k^{-(1+\beta)}$). It is well known that branching laws of this type are in the domain of attraction of the $(1+\beta)$-stable law. Moreover, the limits of fluctuations related to infinite variance branching systems of type $1+\beta$ are usually $(1+\beta)$-stable. (See for example Iscoe [14] Theorem 5.4 and 5.6 and Bojdecki et al. [10]). Another connection is due to [9], relating Beta$(1-\beta,\beta)$-coalescents to continuous state $(1+\beta)$-stable processes. The limit process is naturally totally skewed to the left, as N_t only has negative jumps, hence so does X_ε. □

Define $X^*_\varepsilon(0) = 0$, $X^\beta_\varepsilon(0) = 0$ and

$$X^*_\varepsilon(t) = \varepsilon^{-\frac{1}{1+\beta}} \left(\frac{N_{\varepsilon t}}{v^*_{\varepsilon t}} - 1 \right), \; X^\beta_\varepsilon(t) = \varepsilon^{-\frac{1}{1+\beta}} \left((\varepsilon t)^{1/\beta} \frac{N_{\varepsilon t}}{K_1} - 1 \right), \; t > 0, \tag{4.6}$$

where K_1 is the constant given by (4.1). Let $\Longrightarrow$ denote the convergence in law of processes with respect to the Skorokhod topology.

As a corollary to Theorem 4.2 we obtain the following results.

Theorem 4.4. *Assume* (A), *and let Z and K be as in Theorem 4.2. Then*
(a) $X^*_\varepsilon \Longrightarrow Z$,
(b) if moreover $(y^\beta g(y) - A) = O(y^\alpha)$, *as* $y \to 0$, *for some* $\alpha > \beta/(1+\beta)$, *then*

$$X^\beta_\varepsilon \Longrightarrow Z.$$

Remark 4.5. The proof is postponed until Section 8. As a counterpart to part (b) we exhibit (in Section 8.2) a family of counterexamples, for which $y \mapsto y^\beta g(y)$ is not sufficiently Hölder continuous at 0, and the above "natural extension" of convergence in Theorem 4.2 (b) fails. In turns out that one does not have to search hard for counterexamples: the first guess $g(y) = y^{-\beta} + y^{\alpha-\beta}$, where α is such that $\alpha < \beta/(\beta+1)$, already does the trick. This illustrates a remarkable sensitivity of the second order approximation for N with respect to the smoothness of Λ near 0. □

5 Proofs of the properties of Ψ and v

In this section we prove Lemmas 2.1 and 2.2, and collect some other properties of Ψ and v under assumption (A), which will be used later (Lemma

5.1).

Proof of Lemma 2.1. We start with some useful representations for Ψ. Clearly $\Psi(1) = 0$ and if $q > 1$ we have

$$\Psi(q) =q \int_0^1 \int_0^y \left(1 - (1 - r)^{q-1}\right) dr \frac{\Lambda(dy)}{y^2} \tag{5.1}$$

$$=q(q-1) \int_0^1 \int_0^y \int_0^r (1 - u)^{q-2} du dr \frac{\Lambda(dy)}{y^2} \tag{5.2}$$

$$=q(q-1) \int_0^1 \int_0^1 \int_0^r (1 - uy)^{q-2} du dr \Lambda(dy). \tag{5.3}$$

Representation (5.3) shows that Ψ is finite, continuous on $[1, \infty)$, and strictly positive on $(1, \infty)$. Note that if $q \geq 2$, then the integrand in (5.3) is smaller than 1 so $\psi(q) \leq q(q-1)/2$. The general estimate (2.2) follows from (5.3), the fact that for $0 \leq u, y \leq 1$ and $q \geq 1$ we have $(1-uy)^{q-2} \leq (1-u)^{-1}$ (easy for $q = 1$], and then use monotonicity) and finally the identity $\int_0^1 \log(1 - r)\, dr = -1$. The estimates of type (2.3) were already derived in [3, 17, 16]. The lower bound is a consequence of (1.2), (1.5) and the trivial inequality $(1 - y)^q \leq e^{-qy}$ for $0 \leq y \leq 1$. The upper bound can obtained for example by using (5.1) and its analogue for Ψ^* that yield

$$\Psi^*(q) - \Psi(q) = q \int_0^1 \int_0^y \left((1 - r)^{q-1} - e^{-qr}\right) dr \frac{\Lambda(dy)}{y^2},$$

and observing that $(1-r)^{q-1} - e^{-qr} \leq (1-r)^{q-1} - (1-r)^q \leq r$ for $0 \leq r \leq 1$ and $q \geq 1$. The bound (2.4) follows easily from (5.1), and (2.5) can be proved via a similar representation for Ψ^*. Given (i)–(iii) and the above argument, for (iv) it clearly suffices to show that h is increasing and differentiable. This can be easily seen from (5.1). $\square$

Proof of Lemma 2.2. We have $\Psi(1) = 0$. Moreover, (2.2) shows that $\int_1^\infty dq/\Psi(q) = \infty$. Together with the strict positivity of Ψ on $(1, \infty)$ and (1.6), this implies that $x \to F(x) := \int_x^\infty dq/\Psi(q)$ maps $(1, \infty)$ bijectively to $(0, \infty)$. In particular, v is well defined as the inverse of F, it is clearly a strictly decreasing function and (i) holds. Property (ii) is clear by the definition of v and fundamental theorem of calculus. Provided we show the claim in (iii), (iv) is clearly true due to (1.4). Similarly

$$\frac{N_t}{v_t} - 1 = \frac{v_t^*}{v_t} \left(\frac{N_t}{v_t^*} - 1\right) + \frac{v_t^*}{v_t} - 1,$$

so (iii) and [4] Theorem 2 together imply (2.9). The estimate in (2.10) follows easily from (2.9) by the triangle inequality, the (decreasing) monotonicity of N, and the fact that $v_t \in (1, \infty)$ for each $t > 0$.

In the rest of the argument we prove (iii). This deterministic argument is a simplified version of the stochastic (martingale based) argument for [3], Theorem 1. We will show a somewhat stronger statement that $\log \frac{v_t}{v_t^*} = O(t)$ as $t \to 0+$. In order to do this, for $n \in \mathbb{N}$, $n > 1$ define the functions $v^{(n)}$ and $v^{*,(n)}$ by

$$t = \int_{v_t^{(n)}}^n \frac{1}{\Psi(q)} dq \quad \text{and} \quad t = \int_{v_t^{*,(n)}}^n \frac{1}{\Psi^*(q)} dq.$$

By Lemma 2.1, Ψ is strictly positive on $(1,\infty)$ and it satisfies $\int_1^n \frac{dq}{\Psi(q)} = \infty$, hence $v_t^{(n)}$ is well defined. Similarly, it is easy to see (and checked in [3]) that Ψ^* is strictly positive on $(0,\infty)$ and $\int_0^n \frac{dq}{\Psi^*(q)} = \infty$, so $v_t^{*,(n)}$ is also well defined. Moreover, by (1.3) and (1.6) for each $t>0$ we have that $v_t^{(n)} \nearrow v_t$ and $v_t^{*,(n)} \nearrow v_t^*$ as $n\to\infty$. The functions $v^{(n)}$ and $v^{*,(n)}$ satisfy equations

$$v_t^{(n)} = n - \int_0^t \Psi(v_s^{(n)})ds \quad \text{and} \quad v_t^{*,(n)} = n - \int_0^t \Psi^*(v_s^{*,(n)})ds.$$

Hence $d\log v_t^{(n)} = -\Psi(v_t^{(n)})/v_t^{(n)}\,dt$ and $d\log v_t^{*,(n)} = -\Psi^*(v_t^{*,(n)})/v_t^{*,(n)}\,dt$. This implies that

$$\log\frac{v_t^{(n)}}{v_t^{*,(n)}} + \int_0^t \left[\frac{\Psi(v_s^{(n)})}{v_s^{(n)}} - \frac{\Psi^*(v_s^{*,(n)})}{v_s^{*,(n)}}\right] ds = 0.$$

Observe also that if t is sufficiently small then $v_t^* \geq 2$. Hence there exists a $t_2^*>0$ such that for all sufficiently large n we have $\inf_{t\in[0,t_2^*]} v_t^{*,(n)} > 1$. For such n and t one can rewrite the last identity as

$$\log\frac{v_t^{(n)}}{v_t^{*,(n)}} + \int_0^t \left[\frac{\Psi(v_s^{(n)})}{v_s^{(n)}} - \frac{\Psi(v_s^{*,(n)})}{v_s^{*,(n)}}\right] ds = \int_0^t \frac{\Psi^*(v_s^{*,(n)}) - \Psi(v_s^{*,(n)})}{v_s^{*,(n)}} ds. \tag{5.4}$$

By (2.3) the absolute value of the integral on the right hand side of this equation is bounded by $\frac{t}{2}$. Moreover, by Lemma 2.1 (iv) the function $q \mapsto \Psi(q)/q$ is strictly increasing, so we can apply Lemma 2.4 obtaining $\left|\log(v_t^{(n)}/v_t^{*,(n)})\right| \leq t$. Letting $n\to\infty$ we get $\left|\log\frac{v_t}{v_t^*}\right| \leq t$, thus finishing the proof. □

Under Assumption (A) it is possible to obtain exact asymptotics of Ψ and v, as given by the following lemma.

Lemma 5.1. *Assume* (A). *Then*
(i)

$$\lim_{q\to\infty}\frac{\Psi(q)}{q^{1+\beta}} = \lim_{q\to\infty}\frac{\Psi^*(q)}{q^{1+\beta}} = \frac{A\Gamma(1-\beta)}{\beta(\beta+1)}, \tag{5.5}$$

(ii)

$$\lim_{t\to 0+} t v_t^\beta = \lim_{t\to 0+} t(v_t^*)^\beta = \frac{1+\beta}{A\Gamma(1-\beta)}. \tag{5.6}$$

Moreover, there exist $C_1, C_2>0$ such that for all $t>0$

$$C_1 t^{-\frac{1}{\beta}} \leq v_t \leq C_2\left(t^{-\frac{1}{\beta}} \vee 1\right). \tag{5.7}$$

(iii) For h defined by (2.1) *we have*

$$\lim_{q\to\infty} q^{1-\beta}h'(q) = \frac{A\Gamma(1-\beta)}{1+\beta}, \tag{5.8}$$

moreover

$$\sup_{q\geq 1} q^{1-\beta}h'(q) < \infty. \tag{5.9}$$

Proof. (i) From Assumption (A) it follows that there exists $0 < a < \frac{1}{2}$ such that Λ has a density g on $[0, a]$ and

$$\frac{A}{2} \leq \inf_{0<y\leq a} g(y)y^{\beta} \leq \sup_{0<y\leq a} g(y)y^{\beta} \leq 2A. \tag{5.10}$$

Due to (2.3)–(2.5) it suffices to prove (5.5) with Ψ_a^*. It is immediate to check that $\Psi_a^*(q) = q^2 \int_0^1 \int_0^1 \int_0^r e^{-quy} dudr\Lambda_a(dy)$ (note that this is an analogue of (5.3)). Hence

$$\begin{aligned}
\lim_{q\to\infty} \frac{\Psi_a^*(q)}{q^{1+\beta}} &= \lim_{q\to\infty} q^{1-\beta} \int_0^1 \int_0^r \int_0^a e^{-qyu} g(y) dydudr \\
&= \lim_{q\to\infty} \int_0^1 \int_0^r \int_0^{auq} u^{\beta-1} e^{-y} y^{-\beta} g(\frac{y}{qu})(\frac{y}{qu})^{\beta} dydudr \\
&= \frac{A\Gamma(1-\beta)}{\beta(1+\beta)},
\end{aligned}$$

where the second equality is obtained via the substitution $y' = uqy$ (then y' is renamed y) while the third follows by (A), (5.10) and the dominated convergence theorem.

(ii) Due to (1.7) and the fact that v diverges to ∞ at 0, we have

$$\lim_{t\to 0} tv_t^{\beta} = \lim_{x\to\infty} x^{\beta} \int_x^{\infty} \frac{1}{\Psi(q)} dq,$$

and by the l'Hospital rule and (5.5) we obtain that $\lim_{t\to 0} tv_t^{\beta} = \frac{1+\beta}{A\Gamma(1-\beta)}$. The same is true for v^*. Finally note that (5.7) follows from (5.6), the (decreasing) monotonicity of v, and the fact that $v_t > 1$ for all t.

(iii) Let a be as in the proof of part (i). By (5.1) we have that

$$h = h_a + \tilde{h}_a, \tag{5.11}$$

where

$$h_a(q) = \int_0^a \int_0^y \left(1 - (1-r)^{q-1}\right) dr \frac{\Lambda(dy)}{y^2}, \tag{5.12}$$

$$\tilde{h}_a(q) = \int_a^1 \int_0^y \left(1 - (1-r)^{q-1}\right) dr \frac{\Lambda(dy)}{y^2}. \tag{5.13}$$

Then

$$h_a'(q) = \int_0^a \int_0^y \left(-\ln(1-r)\right) (1-r)^{q-1} dr \frac{g(y)}{y^2} dy \tag{5.14}$$

and

$$\tilde{h}_a'(q) = \int_a^1 \int_0^y \left(-\ln(1-r)\right) (1-r)^{q-1} dr \frac{\Lambda(dy)}{y^2}. \tag{5.15}$$

In the above expression for $\tilde{h}_a'$ we substitute $r' = -\ln(1-r)$ and use the obvious estimates to get

$$\tilde{h}_a'(q) \leq \frac{1}{a^2} \int_0^{\infty} re^{-rq}\, dr = \frac{1}{a^2 q^2}. \tag{5.16}$$

For h'_a we first use the substitution $r' = \frac{r}{y}$ and then $y' = y(q-1)r'$ to obtain

$$q^{1-\beta}h'_a(q) = \frac{q^{1-\beta}}{(q-1)^{1-\beta}} \int_0^1 \int_0^{a(q-1)r} \frac{\left(-\ln(1-\frac{y}{q-1})\right)}{\frac{y}{q-1}} (1-\frac{y}{q-1})^{q-1} \times r^\beta \frac{g(\frac{y}{r(q-1)})(\frac{y}{r(q-1)})^\beta}{y^\beta} dydr. \tag{5.17}$$

Hence again (A), (5.10) and the dominated convergence theorem yield

$$\lim_{q\to\infty} q^{1-\beta}h'_a(q) = \frac{A\Gamma(1-\beta)}{1+\beta}. \tag{5.18}$$

Here we use the facts that $(1-\frac{y}{q-1})^{q-1} \le e^{-y}$, $-\ln(1-z)/z \to 1$ as $z \to 0$, and also that $\sup_{z\le ar<1/2} -\ln(1-z)/z$ is a finite quantity. Now (5.11), (5.16) and (5.18) jointly imply (5.8).

The expression (5.17) and the bounds just used in deriving (5.8) also imply that the function $q \mapsto q^{1-\beta}h'_a(q)$ is bounded on $[2,\infty)$ and, due to the global continuity of h'_a, we conclude that the same function is bounded on $[1,\infty)$. Together with (5.16) and (5.11), this proves (5.9).

6 Complement to Section 3

Proof of Lemma 3.5. Recall the notation introduced at the beginning of Section 3. Let us first notice that $f(1,\cdot,\cdot) \equiv 0$. For $k \in \mathbb{N}$, $k > 0$ it is easy to derive (see also [3], Lemma 17 (iii))

$$\begin{aligned}\int_{[0,1]^{\mathbb{N}}} f^2(k,y,\mathbf{x})d\mathbf{x} &= E[\text{Binomial}(k,y) - \mathbf{1}_{\{\text{Binomial}(k,y)>0\}}]^2 \\ &= k(k-1)y^2 - k(k-1)\int_0^y \int_0^r (1-u)^{k-2}dudr. \end{aligned} \tag{6.1}$$

Hence

$$E\int_0^t \int_{[0,1]\times[0,1]^{\mathbb{N}}} \left(\frac{f(N_{s-},y,\mathbf{x})}{v_s}\right)^2 \frac{\Lambda(dy)}{y^2} dsd\mathbf{x} \le E\int_0^t \int_0^1 \frac{N_{s-}(N_{s-}-1)}{v_s^2}\Lambda(dy)ds \le Ct, \tag{6.2}$$

where the last inequality follows from the second moment estimates in Lemma 2.2 (v), and the continuity of v.

Due to the standard properties of integrals with respect to the compensated Poisson random measure (see e.g. Theorem 8.23 in [18]), (6.2) now implies that $\tilde{M}$ given by (3.8) is well defined square integrable martingale with quadratic variation (3.9). Moreover,

$$E[\tilde{M}](t) = \int_{[0,t]\times[0,1]\times[0,1]^{\mathbb{N}}} E\left(\frac{f(N_{s-},y,\mathbf{x})}{v_s}\right)^2 ds\frac{\Lambda(dy)}{y^2}d\mathbf{x}.$$

Hence (3.10) for $p = 2$ is a consequence of (6.2) and the Doob inequality. The assertion for $0 < p < 2$ then follows due to Jensen's inequality.
Proof of Lemma 3.6. Due to Lemma 2.1 we know that for any $s > 0$, $\frac{N_s}{v_s}\left(\frac{\Psi(N_s)}{N_s} - \frac{\Psi(v_s)}{v_s}\right)$ has the same sign as $\frac{N_s}{v_s} - 1$, hence by Lemmas 3.1–3.5 (after subtracting 1 on both sides of (3.7)) and Lemma 2.4 we obtain

$$\sup_{r \le s \le t} \left| \frac{N_s}{v_s} - 1 \right| \le \left| \frac{N_r}{v_r} - 1 \right| + \left| \tilde{M}_r \right| + \sup_{r \le s \le t} \left| \tilde{M}_s \right|. \tag{6.3}$$

Now Lemma 3.5 (iii) implies

$$E \sup_{r \le s \le t} \left| \frac{N_s}{v_s} - 1 \right|^p \le 3^p \left(E \left| \frac{N_r}{v_r} - 1 \right|^p + E \left| \tilde{M}_r \right|^p + C(p) t^{\frac{p}{2}} \right).$$

Letting $r \to 0$, and using (2.9) and once again (3.10), we obtain (3.11).

7 Proof of Theorem 4.2

We start this section by giving the scheme (or skeleton) of the proof, including an informal (heuristic) discussion on why Theorem 4.2 should hold. Our argument is divided into several lemmas, which are proved separately in the forthcoming subsections.

The first few steps were carried out in Sections 2 and 3 under the only assumption that the coalescent comes down from infinity. Here, as was already done in the final part of Section 5, we specialize further to the case when Λ satisfies Assumption (A). Recall that (A) implies CDI. Throughout this section we assume (A) without much further mention.

The following result is a consequence of Lemmas 3.3, 3.5 and 3.6, where Assumption (A) makes passing to the limit $r \searrow 0$ possible in the identity (3.7).

Proposition 7.1. *We have*

$$\frac{N_t}{v_t} - 1 = -\int_0^t \frac{N_s}{v_s} \left(\frac{\Psi(N_s)}{N_s} - \frac{\Psi(v_s)}{v_s} \right) ds - \tilde{M}_t, \quad t \ge 0, \tag{7.1}$$

almost surely, where $\tilde{M}$ is defined by (3.8)*.*

Remark 7.2. In the general case (without assuming (A)) one can similarly obtain a weaker identity (holding only in L^2), where the L^2 limit

$$\lim_{r \to 0} \int_r^t \frac{N_s}{v_s} \left(\frac{\Psi(N_s)}{N_s} - \frac{\Psi(v_s)}{v_s} \right) ds$$

exists and replaces the integral from 0 to t in (7.1). At the moment we do not know whether $s \mapsto \frac{N_s}{v_s}\left(\frac{\Psi(N_s)}{N_s} - \frac{\Psi(v_s)}{v_s}\right)$ is almost surely Lebesgue integrable on $[0, t]$ in general. $\square$

If $\mathbf{X} = (X_1, X_2, \ldots)$, where X_i, $i = 1, 2, \ldots$ are i.i.d. random variables uniformly distributed on $[0,1]$, then due to the form of f (see (3.3)) and the law of large numbers it is clear that, for each fixed y,

$$\lim_{k\to\infty} \frac{f(k, y, \mathbf{X})}{k} = y \ a.s.$$

Accounting for (2.8) and $\lim_{t\to 0} v_t = \infty$, one would expect that for small t $\tilde{M}$ should be close to a martingale $M = (M(t))_{t\geq 0}$ defined by

$$M(t) = \int_{[0,t]\times[0,1]} y\hat{\pi}(dsdy) \tag{7.2}$$

where $\hat{\pi}$ is the compensated Poisson random measure π (see (3.1)), i.e.

$$\hat{\pi}(dsdy) = \pi(dsdy) - ds\frac{\Lambda(dy)}{y^2}. \tag{7.3}$$

Note that M is a Lévy process with the Lévy measure $\frac{\Lambda(dy)}{y^2}$.

The above heuristic indeed turns out to be true. More precisely, we have the following estimate of the difference of $\tilde{M}$ and M:

Lemma 7.3. *There exist $t_0 > 0$ and $0 < C < \infty$ such that for all $0 < t \leq t_0$*

$$E\sup_{s\leq t}\left(\tilde{M}_s - M_s\right)^2 \leq C(t^2 \vee t^{\frac{1}{\beta}}). \tag{7.4}$$

Concerning the integral on the right hand side of (7.1) we have

Lemma 7.4. *There exist $t_0 > 0$ and $0 < C < \infty$ such that for all $0 < t \leq t_0$*

$$E\sup_{u\leq t}\left|\int_0^u \frac{N_s}{v_s}\left(\frac{\Psi(N_s)}{N_s} - \frac{\Psi(v_s)}{v_s}\right)ds - \int_0^u \left(\frac{N_s}{v_s} - 1\right)v_s h'(v_s)ds\right| \leq Ct, \tag{7.5}$$

where h is defined by (2.1).

Let us denote by X the process

$$X(t) = \frac{N_t}{v_t} - 1,\ t > 0,\ X(0) = 0. \tag{7.6}$$

Then

$$X_\varepsilon = \left(\varepsilon^{-\frac{1}{1+\beta}}X(\varepsilon t),\ t \geq 0\right)$$

is the same as the process X_ε defined in (4.2).

Digression - heuristics. At this point it is possible to explain why the limit process of Theorem 4.2 is of the form as in (4.3) (the longer rigorous argument is given below). From (5.6) and (5.8) it is not difficult to see that for s close to zero we have $v_s h'(v_s) \sim \frac{1}{s}$. Proposition 7.1 and Lemmas 7.3–7.4 then jointly give

$$X(t) \approx -\int_0^t X(s)\frac{1}{s}ds - M_t.$$

Making a change of variables in the drift part we would then have

$$X_\varepsilon(t) \approx -\int_0^t X_\varepsilon(s)s^{-1}ds - M_\varepsilon(t),$$

where

$$M_\varepsilon(t) = \varepsilon^{-\frac{1}{1+\beta}}M(\varepsilon t). \tag{7.7}$$

By investigating the Laplace transform of M_ε it is not difficult to see that it converges in the sense of finite dimensional distributions to KL, where L is the Lévy process described in Remark 4.3 (b) (this can be verified similarly to Lemma 7.7 below). Then it is natural to suspect that, if the limit Z of X_ε exists, it should satisfy the equation given in (4.5). This is indeed the case for the process Z of Theorem 4.2.

There are a few delicate points in the above reasoning. We were unable to replace $v_s h'(v_s)$ directly by $\frac{1}{s}$ and still get a sufficiently good estimate (analogous to that of Lemma 7.4) on the difference of the corresponding integrals. Furthermore, the convergence of X_ε has to be proved, and the passage to the limit under the integral justified.

Our rigorous argument is continued in the following way. Define

$$Y(t) = \int_{[0,t]} \frac{h(v_t)}{h(v_s)}dM(s), \ t \geq 0, \tag{7.8}$$

where as usual h is given by (2.1), and M by (7.2). We will need the following lemma.

Lemma 7.5. *The process Y is the unique solution of the equation*

$$dY(t) = -Y(t)v_t h'(v_t)dt + dM(t), \ Y(0) = 0. \tag{7.9}$$

Next, we prove that the process $-Y$ is close to X.

Lemma 7.6. *There exist $t_0 > 0$ and $C > 0$ such that*

$$E\sup_{u\leq t}|X(u) + Y(u)| \leq C\left(t \vee t^{\frac{1}{2\beta}}\right), \ \forall t \leq t_0. \tag{7.10}$$

Let Y_ε denote the following scaled process

$$Y_\varepsilon(t) = \varepsilon^{-\frac{1}{1+\beta}}Y(\varepsilon t), \ t \geq 0. \tag{7.11}$$

Since $1 > \frac{1}{1+\beta}$ and $\frac{1}{2\beta} > \frac{1}{1+\beta}$ for $0 < \beta < 1$, then Lemma 7.6 implies that $E\sup_{t\leq T}|X_\varepsilon(t) + Y_\varepsilon(t)| \to 0$, for each fixed $T > 0$. In order to prove Theorem 4.2 it therefore suffices to show that, as $\varepsilon \to 0$, Y_ε converges in law to $-Z$ (Z is as defined in (4.3)) with respect to the Skorokhod topology on $D([0,\infty))$, as $\varepsilon \to 0$.

Here we proceed in the standard way: we first derive the convergence of finite dimensional distributions via the Laplace transform, and then prove tightness by means of Aldous' tightness criterion. Let Z be given in (4.3).

Lemma 7.7. *As $\varepsilon \to 0$, Y_ε converges to $-Z$ in the sense of finite dimensional distributions.*

Lemma 7.8. *We have that $Y_\varepsilon \Longrightarrow -Z$ as $\varepsilon \to 0$.*

This final lemma, joint with the discussion following the statement of Lemma 7.6, concludes the proof of Theorem 4.2.

7.1 Proof of Proposition 7.1

Let us subtract 1 on both sides of (3.7) and send $r \to 0$. Lemma 3.5 (more precisely, (3.8) and (3.10)) implies that the integral with respect to $\hat{\pi}^E$ converges in L^2 to $\tilde{M}_t$, while Lemma 2.2 part (v) implies that $\frac{N_r}{v_r} - 1$ converges to 0 in L^2. Therefore, the remaining term on the right hand side of (3.7) must also converge in L^2. Moreover, it is not hard to see that the integral

$$\int_0^t \frac{N_s}{v_s} \left(\frac{\Psi(N_s)}{N_s} - \frac{\Psi(v_s)}{v_s} \right) ds = \int_0^t \frac{N_s}{v_s} (h(N_s) - h(v_s))\, ds$$

is well defined a.s. as a Lebesgue integral. Indeed, the derivative of h is non-negative due to Lemma 2.1 part (iv). We will repeatedly use Assumption (A) in the rest of the argument. Observe that (5.11)–(5.15) imply that h' is decreasing. Hence, if $N_s \le v_s$ then

$$\begin{aligned} \frac{N_s}{v_s} |h(N_s) - h(v_s)| \le & N_s h'(N_s) \left| \frac{N_s}{v_s} - 1 \right| \\ \le & C(\frac{1}{s} \vee 1) \left| \frac{N_s}{v_s} - 1 \right|, \end{aligned}$$

where the last inequality follows from (5.9), the fact that $N_s^\beta \le v_s^\beta$ and (5.7).

If $N_s > v_s$, then, again by (5.7) and (5.9)

$$\begin{aligned} \frac{N_s}{v_s} |h(N_s) - h(v_s)| \le & N_s h'(v_s) \left| \frac{N_s}{v_s} - 1 \right| \\ \le & C(\frac{1}{s} \vee 1) \frac{N_s}{v_s} \left| \frac{N_s}{v_s} - 1 \right|. \end{aligned}$$

The Cauchy-Schwarz inequality, Lemma 3.6 and (2.10) now imply that

$$\begin{aligned} E\left(\int_0^t \frac{N_s}{v_s} \left| \frac{\Psi(N_s)}{N_s} - \frac{\Psi(v_s)}{v_s} \right| ds \right) & \le & CE \int_0^t \left(\frac{1}{s} \vee 1 \right) \left(1 + \frac{N_s}{v_s} \right) \left| \frac{N_s}{v_s} - 1 \right| ds \\ & \le & C_1 \int_0^t \left(\frac{1}{s} \vee 1 \right) \sqrt{s}\, ds < \infty. \end{aligned}$$

Letting $r \to 0$ in (3.7) we obtain (7.1) a.s. for any fixed $t > 0$. The processes on both sides of the equation (7.1) are right continuous, hence they are indistinguishable.

7.2 Proof of Lemma 7.3

Recalling the forms of M and $\tilde{M}$ (see (7.2) and (3.8)) as well as (3.1), observe that $\tilde{M} - M$ is a square integrable martingale with quadratic variation process

$$[\tilde{M} - M](t) = \int_{[0,t]\times[0,1]\times[0,1]^{\mathbb{N}}} \left(\frac{f(N_{s-}, y, \mathbf{x})}{v_s} - y \right)^2 \pi^E(dsdyd\mathbf{x}).$$

Thus we have

$$E[\tilde{M} - M](t) \le 2EI_1(t) + 2EI_2(t),$$

where

$$I_1(t) = \int_{[0,t]\times[0,1]\times[0,1]^{\mathbb{N}}} \left(\frac{f(N_{s-}, y, \mathbf{x}) - N_{s-}y}{v_s} \right)^2 ds \frac{\Lambda(dy)}{y^2} d\mathbf{x} \qquad (7.12)$$

and

$$I_2(t) = \int_{[0,t]\times[0,1]} \left(\frac{N_s}{v_s} - 1 \right)^2 ds\Lambda(dy) = \int_0^t \left(\frac{N_s}{v_s} - 1 \right)^2 ds. \qquad (7.13)$$

By Doob's inequality, it therefore suffices to show

$$EI_i(t) \leq C\left(t^2 \vee t^{\frac{1}{\beta}}\right), \ \ i = 1, 2. \qquad (7.14)$$

Estimate (7.14) for I_2 is immediate by Lemma 3.6. Arguing (7.14) for I_1 is a bit more involved. Let us denote

$$J(k) = \int_0^1 \int_{[0,1]^{\mathbb{N}}} (f(k, y, \mathbf{x}) - ky)^2 \, d\mathbf{x} \frac{\Lambda(dy)}{y^2} \ \ k \in \mathbb{N}, \qquad (7.15)$$

so that

$$I_1(t) = \int_{[0,t]} \frac{J(N_{s-})}{v_s} ds.$$

By (3.3), (6.1) and the following easy to check identity

$$\int_{[0,1]^{\mathbb{N}}} f(k, y, \mathbf{x}) d\mathbf{x} = ky - k \int_0^y (1 - r)^{k-1} dr,$$

we have

$$J(k) \leq 2k^2 \int_0^1 \int_0^y (1 - r)^{k-1} \cdot y \, dr \frac{\Lambda(dy)}{y^2}.$$

Taking a which satisfies (5.10), and applying $1 - r \leq e^{-r}$ we write

$$J(k) \leq 2e\left(J_a(k) + \tilde{J}_a(k)\right), \qquad (7.16)$$

where

$$J_a(k) = k^2 \int_0^a \int_0^y e^{-kr} dr \frac{\Lambda(dy)}{y}, \quad \tilde{J}_a(k) = k^2 \int_a^1 \int_0^y e^{-kr} dr \frac{\Lambda(dy)}{y}.$$

By (5.10) and the natural substitutions ($r' = r/y$, followed by $y' = kr'y$, and afterwards r', y' renamed to r, y, respectively) we have

$$J_a(k) \leq Ck^{1+\beta} \int_0^1 \int_0^{akr} e^{-y} y^{-\beta} r^{\beta-1} dy dr \leq C_1 k^{1+\beta}.$$

The term $\tilde{J}_a$ can be easily bounded as follows

$$\tilde{J}_a(k) \leq \frac{k}{a}.$$

Recalling (7.16), we therefore have $J(k) \leq Ck^{1+\beta}$ for some $C < \infty$. Together with (7.15), (7.12), (2.10) and (5.7), this now implies that (for $t_0 < 1/2$ we use $1 \vee 1/s = 1/s$, $\forall s < t_0$)

$$\begin{aligned} EI_1(t) =& E\int_0^t J(N_{s-})\frac{1}{v_s}\,ds \leq CE\int_0^t \left(\frac{N_s}{v_s}\right)^{1+\beta} v_s^{\beta-1}\,ds \\ \leq& C_1\int_0^t \left(\frac{1}{s^{1/\beta}}\right)^{\beta-1} ds = C_2 t^{1/\beta}, \end{aligned}$$

which proves (7.14) for $i = 1$, and completes the argument.

7.3 Proof of Lemma 7.4

Let h be defined by (2.1) and let h_a and $\tilde{h}_a$ be as in (5.12)–(5.13), with $0 < a < \frac{1}{2}$ satisfying (5.10). Using the easy estimate $\tilde{h}_a(q) \leq a^{-2}$ together with (2.10) we have

$$E\int_0^t \frac{N_s}{v_s}\left|\tilde{h}_a(N_s) - \tilde{h}_a(v_s)\right| ds \leq Ct.$$

Moreover, by (5.16), Lemma 3.6 and (5.7) we obtain

$$E\int_0^t \left|\frac{N_s}{v_s} - 1\right| v_s\tilde{h}'_a(v_s)ds \leq Ct^{\frac{1}{\beta}+\frac{3}{2}}.$$

Hence to prove the Lemma it suffices to show (7.5) with h replaced by h_a. Using the Taylor expansion formula we write

$$\frac{N_s}{v_s}\left(h_a(N_s) - h_a(v_s)\right) = I_1(s) + I_2(s), \tag{7.17}$$

where

$$I_1(s) = \frac{N_s}{v_s}\frac{N_s - v_s}{v_s}v_s h'_a(v_s), \quad I_2(s) = \frac{N_s}{v_s}\int_{v_s}^{N_s}\int_{v_s}^{z} h''_a(w)dwdz.$$

We shall prove that I_1 is the main term, uniformly close to $(N_\cdot - v_\cdot)h'_a(v_\cdot)$, and that I_2 is a negligible error term. First note that by Lemma 3.6, (5.9) (recall that $h'_a \leq h'$) and (5.7) one can easily see that

$$E\left|\left(\frac{N_s}{v_s} - 1\right)(N_s - v_s)h'_s(v_s)\right| \leq CE\left(\frac{N_s}{v_s} - 1\right)^2 v_s^\beta = O(1), \tag{7.18}$$

and therefore

$$E\int_0^t \left|I_1(s) - (N_s - v_s)h'_a(v_s)\right| ds \leq Ct. \tag{7.19}$$

Our approach for I_2 is to show a similar bound

$$|I_2(s)| \leq C\left(\frac{N_s - v_s}{v_s}\right)^2 v_s^\beta, \tag{7.20}$$

and then again use (7.18) to bound $\int_0^t |I_2(s)|\,ds$. First note that from differentiating in (5.14) it follows that h''_a is negative and increasing (its absolute

value is decreasing). Moreover, since $a < \frac{1}{2}$, and since $|\log(1-r)| \le 2r$ and $(1-r)^{q-1} \le 2e^{-rq}$ for $r \le 1/2$, one can easily derive from (5.10) that

$$\left|h_a''(q)\right| \le C \int_0^a \int_0^y r^2 e^{-rq} y^{-2-\beta} dr dy = O\left(q^{\beta-2}\right). \tag{7.21}$$

Thus, if $\frac{1}{2}v_s \le N_s \le 2v_s$ then

$$|h_a''(w)| \le \left|h_a''\left(\frac{1}{2}v_s\right)\right| = O\left(v_s^{\beta-2}\right)$$

and $|I_2(s)| = \frac{N_s}{v_s}(N_s - v_s)^2 O(v_s^{\beta-2})$. Since $N_s/v_s \le 2$, we conclude that (7.20) holds in this case.

If $v_s > 2N_s$ then note that $\int_{v_s}^{N_s}\int_{v_s}^{z} w^{\beta-2}\, dw\, dz =$

$$\begin{aligned}\int_{N_s}^{v_s}\int_{z}^{v_s} w^{\beta-2}\, dw\, dz &\le \frac{1}{1-\beta}\int_{N_s}^{v_s} z^{\beta-1}\, dz\\ &\le \frac{1}{1-\beta}(v_s - N_s)N_s^{\beta-1}.\end{aligned}$$

Hence, by (7.21) and the definition of I_2

$$|I_2(s)| \le C(\frac{v_s - N_s}{v_s})N_s^{\beta}.$$

We also have $N_s^{\beta} \le v_s^{\beta}$ and $1 < 2\frac{v_s - N_s}{v_s}$, so (7.20) follows.

If $2v_s < N_s$ then,

$$\begin{aligned}\frac{N_s}{v_s}\int_{v_s}^{N_s}\int_{v_s}^{z} w^{\beta-2} dw dz \le& C\frac{N_s}{v_s}(N_s - v_s)v_s^{\beta-1}\\ \le& C\left(\frac{N_s}{v_s} - 1\right)^2 v_s^{\beta} + C\left(\frac{N_s}{v_s} - 1\right)v_s^{\beta}.\end{aligned}$$

Together with (7.21) and the definition of $I_2(s)$ this again implies (7.20), since for $2v_s < N_s$ we have $1 < \frac{N_s}{v_s} - 1 < \left(\frac{N_s}{v_s} - 1\right)^2$.

This gives (7.20), and due to the final estimate in (7.18) we get $E\int_0^t |I_2(s)|\, ds \le Ct$, which combined with (7.19) yields(7.5) for h_a. As already argued, this completes the proof of the lemma.

7.4 Proof of Lemma 7.5

Let us first observe that the function $u \mapsto h(v_u)$ defined in (2.1) is positive on $(0, \infty)$ and strictly decreasing, since h is positive and strictly increasing and v is strictly decreasing (see Lemmas 2.1 and 2.2). Moreover, by (5.5) and (5.6) we have that

$$\lim_{u \to 0} uh(v_u) = \frac{1}{\beta}, \tag{7.22}$$

so, there exists t_0 such that

$$\frac{\beta}{2}u \le \frac{1}{h(v_u)} \le 2\beta u,\ 0 < u \le t_0. \tag{7.23}$$

Hence the process Y from (7.8) is well defined. Moreover,

$$E\,(Y(t))^2 = (h(v_t))^2 \int_0^t \int_0^1 \left(\frac{y}{h(v_u)}\right)^2 \frac{\Lambda(dy)}{y^2} \le t, \tag{7.24}$$

since $h(v_t) \le h(v_u)$ for $u \le t$.

The function $u \mapsto h(v_u)$ is clearly continuous and of finite variation on any interval $[r,t]$, $0 < r < t$. A simple integration by parts in (7.8) (written formally as $fg = \int f'g + \int fg'$ with $f(\cdot) = h(v.)$ and $g(\cdot) = \int_0^{\cdot} \frac{1}{h(v_s)} dM_s$), together with the fact that

$$\frac{v_s'}{h(v_s)} = -v_s,$$

(cf. (2.1) and (2.6)) shows that for $0 < r < t$

$$Y_t = Y_r - \int_r^t Y_s v_s h'(v_s) ds + M_t - M_r. \tag{7.25}$$

We now let $r \to 0$ and observe that $M_r \to 0$ a.s. and in L^2, since $E[M](r) = \int_0^r \int_0^1 y^2 \frac{\Lambda(dy)}{y^2} = r$, and $Y_r \to 0$ in L^2 by (7.24). To deal with the remaining term in (7.25) we note that by (5.9) and (5.7) we have

$$0 \le v_s h'(v_s) \le C\left(s^{-1} \vee 1\right).$$

Hence, by (7.24) and Jensen's inequality

$$E \int_0^r \left|Y_s v_s h'(v_s)\right| ds \le \int_0^r \sqrt{s}\left(\frac{1}{s} \vee 1\right) ds \le C(\sqrt{r} \vee r^{3/2}),$$

converges to 0 as $r \to 0$. After sending $r \to 0$ in (7.25) one concludes that Y given by (7.8) satisfies the equation (7.9).

Showing uniqueness is easier. Indeed, if Y_1 and Y_2 are two solutions of (7.9) then

$$Y_1(t) - Y_2(t) = -\int_0^t \left(Y_1(s) - Y_2(s)\right) v_s h'(v_s) ds.$$

Since $v_s h'(v_s)$ is positive (see Lemma 2.1 (iv)), an application of Lemma 2.4 implies $Y_1 - Y_2 \equiv 0$.

7.5 Proof of Lemma 7.6

Recall (7.6) and the corresponding rescaled process X_ε from (4.2). Due to Proposition 7.1 and Lemmas 7.3, 7.4 and 7.5 we obtain

$$X(t) + Y(t) = -\int_0^t (X(s) + Y(s)) v_s h'(v_s) ds + R(t),$$

where R is a process such that for $0 \le t \le t_0$

$$E \sup_{s \le t} |R(s)| \le C\left(t \vee t^{\frac{1}{2\beta}}\right).$$

Since $v_s h'(v_s)$ is positive, another application of Lemma 2.4 completes the proof.

7.6 Proof of Lemma 7.7

The argument relies on convergence of the Laplace transform for positive arguments. Fix $n \in \mathbb{N}$ and $z_j \geq 0$, $t_j > 0$, $j = 1, 2, \ldots n$ and denote

$$F(u) = \sum_{j=1}^{n} z_j \frac{u}{t_j} \mathbb{1}_{[0,t_j]}(u). \tag{7.26}$$

We will show that

$$\lim_{\varepsilon \to 0} E \exp\left\{ -\sum_{j=1}^{n} z_j Y_\varepsilon(t_j) \right\} = \exp\left\{ A \int_0^\infty (e^{-y} - 1 + y) \frac{1}{y^{2+\beta}} dy \int_0^\infty (F(u))^{1+\beta} du \right\}. \tag{7.27}$$

Due to Propositions 3.4.1 and 1.2.12 and (3.4.4) in [21], the right hand side is precisely $E \exp\left\{ -\sum_{j=1}^{n} z_j(-Z(t_j)) \right\}$, where Z is defined in (4.3). On the other hand, it is well known, that since $-Z$ is a $(1+\beta)$-stable process totally skewed to the right, the convergence of Laplace transforms for all positive z_j implies the convergence in law of $(Y_\varepsilon(t_1), \ldots, Y_\varepsilon(t_n))$ to $(-Z(t_1), \ldots, -Z(t_n))$ (see e.g. [14], proofs of Theorems 5.4 and 5.6). Thus the lemma will be proved once we show (7.27).

By (7.8) and (7.11) we have

$$\begin{aligned}\sum_{j=1}^{n} z_j Y_\varepsilon(t_j) &= \varepsilon^{-\frac{1}{1+\beta}} \int_0^\infty \int_0^1 \left(\sum_{j=1}^{n} z_j \mathbb{1}_{[0,\varepsilon t_j]}(u) \frac{h(v_{\varepsilon t_j})}{h(v_u)} \right) y \hat{\pi}(du dy) \\ &= \varepsilon^{-\frac{1}{1+\beta}} \int_0^\infty \int_0^1 F_\varepsilon(\frac{u}{\varepsilon}) y \hat{\pi}(du dy),\end{aligned}$$

where

$$F_\varepsilon(u) = \sum_{j=1}^{n} z_j \frac{h(v_{\varepsilon t_j})}{h(v_{\varepsilon u})} \mathbb{1}_{[0,t_j]}(u). \tag{7.28}$$

Thus, by the usual properties of a Poisson random measure we have

$$E \exp\left\{ -\sum_{j=1}^{n} z_j Y_\varepsilon(t_j) \right\} = e^{I(\varepsilon)}, \tag{7.29}$$

where

$$I(\varepsilon) = \int_0^\infty \int_0^1 \left(e^{-\varepsilon^{-\frac{1}{1+\beta}} F_\varepsilon(\frac{u}{\varepsilon}) y} - 1 + \varepsilon^{-\frac{1}{1+\beta}} F_\varepsilon\left(\frac{u}{\varepsilon}\right) y \right) \frac{\Lambda(dy)}{y^2} du. \tag{7.30}$$

As before, let $0 < a < \frac{1}{2}$ be such that (5.10) holds and write

$$I(\varepsilon) = I_a(\varepsilon) + \tilde{I}_a(\varepsilon), \tag{7.31}$$

where

$$I_a(\varepsilon) = \int_0^\infty \int_0^a \ldots \text{ and } \tilde{I}_a(\varepsilon) = \int_0^\infty \int_a^1 \ldots, \tag{7.32}$$

and the ... above denote the expression under the integral in (7.30). Let us initially consider the term $\tilde{I}_a$. We have

$$\begin{aligned}0 \le \tilde{I}_a(\varepsilon) \le& \int_0^\infty \int_a^1 \varepsilon^{-\frac{1}{1+\beta}} F_\varepsilon\left(\frac{u}{\varepsilon}\right) \frac{\Lambda(dy)}{y}\, du \\ \le& \frac{1}{a}\varepsilon^{1-\frac{1}{1+\beta}} \int_0^\infty F_\varepsilon(u)du.\end{aligned}$$

Recall (7.28) and note that since $h(v_{\varepsilon t}) \le h(v_{\varepsilon u})$ for $u \le t$, as explained in the proof of Lemma 7.5, then $\sup_{\varepsilon>0} \int_0^\infty F_\varepsilon(u)du < \infty$. It follows that

$$\lim_{\varepsilon\to 0} \tilde{I}_a(\varepsilon) = 0. \tag{7.33}$$

In the analysis of $I_a(\varepsilon)$ we make a change of variables $y = z\varepsilon^{1/(1+\beta)}$ and $r = \frac{u}{\varepsilon}$ (then rename z to be y and r to be u) and use Assumption (A) to get

$$I_a(\varepsilon) = \int_0^\infty \int_0^{a\varepsilon^{-\frac{1}{1+\beta}}} \left(e^{-F_\varepsilon(u)y} - 1 + F_\varepsilon(u)y\right) \frac{g(y\varepsilon^{\frac{1}{1+\beta}})\left(y\varepsilon^{\frac{1}{1+\beta}}\right)^\beta}{y^{2+\beta}}\, dy\, du. \tag{7.34}$$

By (5.5), (5.6) and the divergence of $v.$ near 0 we have

$$\lim_{\varepsilon\to 0} \frac{h(v_{\varepsilon t})}{h(v_{\varepsilon u})} = \frac{u}{t},$$

so from (7.28) we see that F_ε converges pointwise to F defined in (7.26). Moreover, note that

$$0 \le e^{-F_\varepsilon(u)y} - 1 + F_\varepsilon(u)y \le F_\varepsilon^2(u)y^2 \le \left(\sum_{j=1}^n z_j \mathbb{1}_{[0,t_j]}(u)\right)^2 y^2.$$

Hence, by (7.34), (A), (5.10) and the dominated convergence theorem it follows that

$$\begin{aligned}\lim_{\varepsilon\to 0} I_a(\varepsilon) =& A\int_0^\infty \int_0^\infty \left(e^{-F(u)y} - 1 + F(u)y\right) \frac{1}{y^{2+\beta}} dy du \\ =& A\int_0^\infty (F(u))^{1+\beta}\, du \int_0^\infty \left(e^{-y} - 1 + y\right) \frac{1}{y^{2+\beta}} dy,\end{aligned} \tag{7.35}$$

where we apply the substitution $z = F(u)y$ and then rename z as y. Now (7.29)–(7.32), (7.33) and (7.35) together imply that (7.27) holds and the proof is complete. □

Proof of Lemma 7.8. First observe that by (7.22) and (7.23) the function f_ε defined by $f(0) = \frac{1}{\beta}$ and $(t) = \varepsilon t h(v_\varepsilon t)$ for $t > 0$ is continuous for any $\varepsilon > 0$. Furthermore, as $\varepsilon \to 0$, the family $(f_\varepsilon)_{\varepsilon>0}$ converge uniformly on bounded intervals to a constant function $\frac{1}{\beta}$. Hence, to prove the lemma it suffices to show that the processes $\tilde{Y}_\varepsilon$ defined by

$$\tilde{Y}_\varepsilon(t) = t^{-1}\beta^{-1}\varepsilon^{-1-\frac{1}{1+\beta}} \int_0^{\varepsilon t} \frac{1}{h(v_u)} dM_u \tag{7.36}$$

converges in law in $D([0,\infty))$ to $-Z$, as $\varepsilon \to 0$.

We will split the proof into several steps. In the first step, with the help of Aldous' tightness criterion, we show that the family of processes $(t\tilde{Y}_\varepsilon(t))_{t\geq 0}$ converges in law in $D([0,\infty))$ to $(-tZ(t))_{t\geq 0}$. From this we need to infer the convergence $\tilde{Y}_\varepsilon \Rightarrow -Z$. However, the latter step is not immediate, since the function $t \mapsto \frac{1}{t}$ cannot be extended to a continuous function on $[0,\infty)$. We will overcome this problem by taking suitable approximations.

Step 1. We prove that the family of processes $(U_\varepsilon)_{\varepsilon>0}$ defined by

$$U_\varepsilon(t) = t\tilde{Y}_\varepsilon(t),\ t \geq 0,\ \varepsilon > 0, \tag{7.37}$$

converges to $(-tZ(t))_{t\geq 0}$ in law in $D([0,\infty))$. It is clearly enough to show this convergence when restricted to an arbitrary but fixed sequence $\varepsilon_n \searrow 0$.

The convergence of finite dimensional distributions follows from (7.22) and Lemma 7.7. To prove tightness of the family $(U_\varepsilon)_{\varepsilon>0}$ we will apply the well-known Aldous criterion (see e.g. [8] Theorem 16.10). More precisely, we will prove:
(i) For any $M > 0$

$$\lim_{r\to\infty} \limsup_{n\to\infty} \mathbb{P}(\sup_{t\in[0,M]} |U_{\varepsilon_n}(t)| \geq r) = 0, \tag{7.38}$$

(ii) For any $\rho, \eta, M > 0$ there exist δ_0, n_0 such that if $\delta \leq \delta_0$, $n \geq n_0$ and τ is a stopping time with respect to the filtration generated by U_{ε_n}, taking finite number of values, and such that $\mathbb{P}(\tau \leq M) = 1$, then

$$\mathbb{P}(|U_{\varepsilon_n}(\tau+\delta) - U_{\varepsilon_n}(\tau)| \geq \rho) \leq \eta. \tag{7.39}$$

To prove (i) and (ii) above, we will need an estimate on the moments of increments of U_ε. We write

$$U_\varepsilon = \frac{1}{\beta}\left(U_\varepsilon^{(1)} + U_\varepsilon^{(2)}\right), \tag{7.40}$$

where

$$\begin{aligned} U_\varepsilon^{(1)}(r) =& \varepsilon^{-\frac{2+\beta}{1+\beta}} \int_{[0,\varepsilon r]\times[0,\varepsilon^{\frac{1}{1+\beta}}]} \frac{1}{h(v_u)} y\hat{\pi}(dudy), \\ U_\varepsilon^{(2)}(r) =& \varepsilon^{-\frac{2+\beta}{1+\beta}} \int_{[0,\varepsilon r]\times(\varepsilon^{\frac{1}{1+\beta}},1]} \frac{1}{h(v_u)} y\hat{\pi}(dudy). \end{aligned}$$

Note that $U_\varepsilon^{(1)}$ (resp. $U_\varepsilon^{(2)}$) is the process which captures the "small" (resp. "large") jumps of U_ε.

Using standard properties of integrals with respect to a compensated Poisson random measure (see e.g. [18], Theorem 8.23) we have

$$E\left|U_\varepsilon^{(2)}(t) - U_\varepsilon^{(2)}(s)\right|^p \leq C\varepsilon^{-\frac{p(2+\beta)}{1+\beta}} \int_{\varepsilon s}^{\varepsilon t} \int_{\varepsilon^{\frac{1}{1+\beta}}}^1 \frac{y^p}{(h(v_u))^p} \frac{\Lambda(dy)}{y^2} du.$$

Let $0 < s < t < T$ and $1 < p < 1+\beta$ and suppose that $\varepsilon \le a^{1+\beta} \wedge \frac{t_0}{T}$, where a is as in (5.10) and t_0 as in (7.23). By (5.10) and (7.23) we obtain

$$\begin{aligned} &E\left|U_\varepsilon^{(2)}(t) - U_\varepsilon^{(2)}(s)\right|^p \\ &\le C\varepsilon^{-\frac{p(2+\beta)}{1+\beta}} \int_{\varepsilon s}^{\varepsilon t} u^p \left(\int_{\varepsilon^{\frac{1}{1+\beta}}}^{a} y^{p-2-\beta} dy + \int_a^1 y^{p-2}\Lambda(dy) \right) du \\ &\le C_1(p) T^p (t-s), \end{aligned} \tag{7.41}$$

since $\varepsilon^{p+1} << \varepsilon^{p+1}\varepsilon^{\frac{p-1-\beta}{1+\beta}} = \varepsilon^{\frac{p(2+\beta)}{1+\beta}}$ cancels the power of ε in front of the integral, and since $\int_a^1 y^{p-2}\Lambda(dy)$ is a constant quantity.

Via similar arguments applied to $U^{(1)}$ we get

$$E\left|U_\varepsilon^{(1)}(t) - U_\varepsilon^{(1)}(s)\right|^2 = \varepsilon^{-\frac{2(2+\beta)}{1+\beta}} \int_{\varepsilon s}^{\varepsilon t} \int_0^{\varepsilon^{\frac{1}{1+\beta}}} \frac{1}{(h(v_u))^2}\Lambda(dy)du,$$

and, since $3 + \frac{1-\beta}{1+\beta} = \frac{2(2+\beta)}{1+\beta}$, again (5.10) and (7.23) yield

$$\begin{aligned} E\left|U_\varepsilon^{(1)}(t) - U_\varepsilon^{(1)}(s)\right|^2 &\le C\varepsilon^{-\frac{2(2+\beta)}{1+\beta}} \int_{\varepsilon s}^{\varepsilon t} \int_0^{\varepsilon^{\frac{1}{1+\beta}}} \frac{u^2}{y^\beta} dy du \\ &\le C_2 T^2 (t-s). \end{aligned} \tag{7.42}$$

Now (7.40)–(7.42) and Jensen's inequality imply that for $0 < s < t < T$ and $1 < p < 1+\beta$, $\varepsilon \le a^{1+\beta} \wedge \frac{t_0}{T}$ we have

$$E\left|U_\varepsilon(t) - U_\varepsilon(s)\right|^p \le C(p) T^p \left(|t-s|^{\frac{p}{2}} \vee |t-s| \right). \tag{7.43}$$

Applying the Doob maximal inequality for martingale U_ε we conclude

$$\mathbb{P}\big(\sup_{t\in[0,M]} |U_\varepsilon(t)| > r\big) \le \left(\frac{p}{p-1}\right)^p \frac{E\left|U_\varepsilon(M)\right|^p}{r^p}.$$

Hence (7.43) implies (7.38).

Estimate (7.43) and the Markov property (since τ takes only finitely many values, we do not need the strong Markov property) of U_ε imply that if τ is a stopping time with respect to the filtration of U_ε taking finite number of values and such that $\tau \le M$, then

$$\begin{aligned} E\left|U_\varepsilon(\tau+\delta) - U_\varepsilon(\tau)\right|^p =& EE\left(\left|U_\varepsilon(\tau+\delta) - U_\varepsilon(\tau)\right|^p |\mathscr{F}_\tau^{U_\varepsilon} \right) \\ \le& C(M+\delta)^p (\delta \vee \delta^{\frac{p}{2}}), \end{aligned}$$

whenever $1 < p < 1+\beta$ and $\varepsilon \le a^{1+\beta} \wedge \frac{t_0}{M+\delta}$. This and the Markov inequality show that condition (ii) (or equivalently, (7.39)) is also satisfied.

As already indicated, using Aldous' criterion we obtain the tightness of the family $(U_{\varepsilon_n})_{n\ge1}$, which together with the already proved convergence of finite dimensional distributions implies that $(U_{\varepsilon_n})_n$ converges in law to $(-tZ(t),\, t \ge 0)$ with respect to the Skorokhod topology on $D([0,\infty))$.

Step 2. For $b > 0$ define

$$Z_\varepsilon^{(b)}(t) = \left(\frac{1}{b} \mathbb{1}_{[0,b]}(t) + \frac{1}{t} \mathbb{1}_{(b,\infty)}(t) \right) U_\varepsilon(t). \tag{7.44}$$

Recall that if $f : \mathbb{R}_+ \mapsto \mathbb{R}$ is continuous then the mapping $w \mapsto fw$ is continuous from $D([0,\infty))$ into itself. Hence the result of Step 1 implies that for any $b > 0$, as $\varepsilon \to 0$, the family of processes $(Z_\varepsilon^{(b)})_{\varepsilon>0}$ converges in law to the process $Z^{(b)}$ defined by

$$Z^{(b)}(t) = \frac{t}{b} \mathbb{1}_{[0,b]}(t) Z(t) + \mathbb{1}_{(b,\infty)}(t) Z(t),\ t \geq 0,$$

with respect to the Skorokhod topology on $D([0,\infty))$.

Step 3. We will next estimate the supremum norms of the difference of $\tilde{Y}_\varepsilon$ and $Z_\varepsilon^{(b)}$, and the difference of Z and $Z^{(b)}$, respectively. Fix any $1 < p < 1+\beta$ and suppose that $b \leq t_0 \wedge 1$ and $\varepsilon \leq a^{1+\beta}$, where t_0 is as in (7.23) and a as in (5.10). Denote $\|f\|_\infty = \sup_{t\in\mathbb{R}_+} |f(t)|$.

Using (7.36)–(7.37) and (7.44) we have that $\tilde{Y}_\varepsilon(t) - Z_\varepsilon^{(b)}(t) = U_\varepsilon(t)(\frac{1}{t} - \frac{1}{b})\mathbf{1}_{\{t\in[0,\frac{1}{b}]\}}$. Therefore,

$$\left\| \tilde{Y}_\varepsilon - Z_\varepsilon^{(b)} \right\|_\infty \leq \sup_{0\leq t\leq b} \left| \tilde{Y}_\varepsilon(t) \right| \leq \sup_{0\leq t\leq b} |Y_\varepsilon(t)|\,,$$

where (7.23) was used in the final estimate. Lemmae 7.5 and 2.4 imply

$$\sup_{0\leq t\leq b} |Y_\varepsilon(t)| \leq \varepsilon^{-\frac{1}{1+\beta}} \sup_{0\leq t\leq b} |M(\varepsilon t)|\,.$$

Hence, decomposing M similarly as it was done for U_ε in Step 1 and applying Doob's inequality for M, we obtain

$$E \left\| \tilde{Y}_\varepsilon - Z_\varepsilon^{(b)} \right\|_\infty^p \leq C_1(p) \left(E \left| \varepsilon^{-\frac{1}{1+\beta}} M^{(1)}(\varepsilon b) \right|^p + E \left| \varepsilon^{-\frac{1}{1+\beta}} M^{(2)}(\varepsilon b) \right|^p \right), \tag{7.45}$$

where

$$M^{(1)}(\varepsilon b) = \int_0^{\varepsilon b} \int_{[0,\varepsilon b]\times[0,\varepsilon^{\frac{1}{1+\beta}}]} y \hat{\pi}(du dy),$$
$$M^{(2)}(\varepsilon b) = \int_{[0,\varepsilon b]\times(\varepsilon^{\frac{1}{1+\beta}},1]} y \hat{\pi}(du dy).$$

By mimicking the arguments of Step 1 we obtain

$$E \left| \varepsilon^{-\frac{1}{1+\beta}} M_{\varepsilon b}^{(1)} \right|^2 \leq C \varepsilon^{-\frac{2}{1+\beta}} \int_0^{\varepsilon b} \int_0^{\varepsilon^{\frac{1}{1+\beta}}} y^{-\beta} dy du = C_1(p) b,$$

and, relying on $\varepsilon << \varepsilon\varepsilon^{\frac{p-1-\beta}{1+\beta}} = \varepsilon^{\frac{p}{1+\beta}}$, we also obtain

$$\begin{aligned} E \left| \varepsilon^{-\frac{1}{1+\beta}} M_{\varepsilon b}^{(1)} \right|^p \leq & C_2(p) \varepsilon^{-\frac{p}{1+\beta}} \int_0^{b\varepsilon} \left(\int_{\varepsilon^{\frac{1}{1+\beta}}}^a y^{p-2-\beta} dy + \int_a^1 y^{p-2} \Lambda(dy) \right) du \\ \leq & C_3(p) b. \end{aligned}$$

Together with (7.45) and Jensen's inequality, for $0 < p < 1+\beta$, $b \le t_0 \wedge 1$ and $\varepsilon \le a^{1+\beta}$, this implies

$$E \left\| \tilde{Y}_\varepsilon - Z_\varepsilon^{(b)} \right\|_\infty^p \le C(p) b^{\frac{p}{2}}, \tag{7.46}$$

where $C(p)$ is some finite constant, uniform in ε.

For the processes $Z^{(b)}$ and Z we again have

$$E \left\| Z - Z^{(b)} \right\|_\infty \le \sup_{t \le b} |Z(t)|.$$

Since Z is a solution of (4.5), we can again apply Lemma 2.4 and Doob's inequality to L, a $(1+\beta)$-stable Lévy process, to derive

$$E \left\| Z - Z^{(b)} \right\|_\infty^p \le C_1(p) E |L(b)|^p \le C_2(p) b^{\frac{p}{1+\beta}}, \tag{7.47}$$

for some $C_2(p) < \infty$.

Step 4. Finally we prove the convergence $\tilde{Y}_\varepsilon \Longrightarrow -Z$ as $\varepsilon \to 0$. Let d_∞^0 denote the Skorokhod metric on $D([0,\infty))$ as defined in [8], p. 168. It is clear that $d_\infty^0(f,g) \le \|f-g\|_\infty$ for any two $f, g \in D([0,\infty))$.

It suffices to show that, whenever $F : D([0,\infty)) \mapsto D([0,\infty))$ is a given bounded and uniformly continuous function, we have

$$\lim_{\varepsilon \to 0} \left| EF(\tilde{Y}_\varepsilon) - EF(Z) \right| = 0. \tag{7.48}$$

By the conclusion of Step 2, for any $b > 0$ we have $E|F(Z_\varepsilon^{(b)}) - EF(Z^{(b)})| \to 0$. Hence (7.48) follows by the triangle inequality, the uniform continuity of F, estimates (7.46) and (7.47) and the Markov inequality, and the above discussion. The argument based on addition and subtraction of intermediate terms is standard, and the details are left to the reader.

8 On robustness with respect to the choice of speed

8.1 Proof of Theorem 4.4

Recall Ψ, Ψ^* and v defined in (1.5), (1.2) and (1.7), respectively. Furthermore recall that v^* is defined in terms of Ψ^* as v is defined in terms of Ψ. Due to (2.3), and in particular to the proof of Lemma 2.2 (iii), one can easily see that

$$\sup_{t \in [0,T]} \frac{1}{\varepsilon^{\frac{1}{1+\beta}}} \left| \frac{v_{\varepsilon t}}{v_{\varepsilon t^*}} - 1 \right| = O(\varepsilon^{1-1/(\beta+1)}), \quad \text{as } \varepsilon \to 0.$$

Since

$$\frac{1}{\varepsilon^{\frac{1}{1+\beta}}} \left(\frac{N_{\varepsilon t}}{v_{\varepsilon t}^*} - 1 \right) = \frac{1}{\varepsilon^{\frac{1}{1+\beta}}} \left(\frac{N_{\varepsilon t}}{v_{\varepsilon t}} - 1 \right) \times \frac{v_{\varepsilon t}}{v_{\varepsilon t}^*} + \frac{1}{\varepsilon^{\frac{1}{1+\beta}}} \left(\frac{v_{\varepsilon t}}{v_{\varepsilon t}^*} - 1 \right), \tag{8.1}$$

one can conclude Theorem 4.4 (a) directly from Theorem 4.2 and (2.7).

We now turn to the proof of part (b). Let us denote $w_t = K_1 t^{-\frac{1}{\beta}}$ for K_1 from (4.1). Observe that an analogue of (8.1), with v^* replaced by w, implies that it suffices to show

$$\lim_{t\to 0} t^{-\frac{1}{1+\beta}} \left(\frac{v_t}{w_t} - 1 \right) = 0. \tag{8.2}$$

Also note that w is related to $\Psi^{(\beta)}(q) = \frac{A\Gamma(1-\beta)}{\beta(1+\beta)} q^{1+\beta}$ via relation

$$t = \int_{w_t} \frac{1}{\Psi^{(\beta)}(q)} dq,$$

the same way that v is related to Ψ (see (1.7)). Recall that from (5.5) we already know $\lim_{q\to\infty} \Psi(q)/\Psi^{(\beta)}(q) = 1$. We will need a more precise comparison of Ψ and $\Psi^{(\beta)}$.

Let $a \le \frac{1}{2}$ be such that Λ has a density g on $[0,a]$ satisfying (5.10) and, moreover $\left| y^\beta g(y) - A \right| \le C y^\alpha$ on $[0,a]$. Such a exists by the assumptions.

Observe that (similarly to derivation of (5.5))

$$\Psi^{(\beta)}(q) = A q^2 \int_0^1 \int_0^r \int_0^\infty e^{-qyu} y^{-\beta} dy du dr. \tag{8.3}$$

Therefore, by Lemma 2.1 (ii) and (iii) we have

$$\Psi(q) = \Psi^*_a(q) + O(q) = \Psi^{(\beta)}(q) + R_1(q) - R_2(q) + O(q), \quad q \ge 1, \tag{8.4}$$

where

$$R_1(q) = q^2 \int_0^1 \int_0^r \int_0^a e^{-qyu} \left(g(y) - A y^{-\beta} \right) dy du dr \tag{8.5}$$

and

$$R_2(q) = q^2 A \int_0^1 \int_0^r \int_a^\infty y^{-\beta} e^{-qyu} dy du dr. \tag{8.6}$$

Due to the assumptions, we have

$$|R_1(q)| \le C q^2 \int_0^1 \int_0^r \int_0^a y^{\alpha-\beta} e^{-qyu} dy du dr. \tag{8.7}$$

If $\alpha < \beta$, then (this is simpler than the proof of Lemma 5.1)

$$|R_1(q)| \le C\Gamma(1+\alpha-\beta) q^{1+\beta-\alpha} \int_0^1 \int_0^r u^{\beta-\alpha-1} du dr = O(q^{1+\beta-\alpha}).$$

If $\alpha \ge \beta$, then by (8.7) we have

$$\begin{aligned} |R_1(q)| &\le C q^2 a^{\alpha-\beta} \int_0^1 \int_0^a e^{-qyu} dy du \\ &\le C a^{\alpha-\beta} \left(q^2 \int_0^{\frac{1}{q}} a du + q \int_{\frac{1}{q}}^1 \frac{1 - e^{-qau}}{u} du \right) \\ &\le C a^{\alpha-\beta} \left(aq + q \int_{\frac{1}{q}}^1 \frac{1}{u} du \right) = O(q(\log q + 1)). \end{aligned}$$

For R_2 we have

$$R_2(q) \le Aq^2 \int_a^\infty y^{-\beta} \int_0^1 e^{-qyu} du dy \le Aq \int_a^\infty y^{-\beta-1} dy = O(q). \tag{8.8}$$

Hence from (8.4) it follows that

$$\Psi(q) = \Psi^{(\beta)}(q) + O(q^{1+\beta-\alpha}) + O(q(\log q + 1)). \tag{8.9}$$

To prove (8.2) we adapt the technique of Lemma 2.2 (iii). In particular, let us consider $v^{(n)}$ and $w^{(n)}$ defined by

$$t = \int_{v_t^{(n)}}^n \frac{1}{\Psi(q)} dq \quad \text{and} \quad t = \int_{w_t^{(n)}}^n \frac{1}{\Psi^{(\beta)}(q)} dq,$$

and the following analogue of (5.4):

$$\log \frac{w_t^{(n)}}{v_t^{(n)}} + \int_0^t \left[\frac{\Psi^{(\beta)}(w_s^{(n)})}{w_s^{(n)}} - \frac{\Psi^{(\beta)}(v_s^{(n)})}{v_s^{(n)}} \right] ds = \int_0^t \frac{\Psi(v_s^{(n)}) - \Psi^{(\beta)}(v_s^{(n)})}{v_s^{(n)}} ds \tag{8.10}$$

(note that if $n \ge 2$ and t is sufficiently small then $w_s^{(n)} \ge 1$ for $s \le t$). Also observe that $v_s^{(n)} \nearrow v_s$, $w_s^{(n)} \nearrow w_s$ as $n \to \infty$. Lemma 2.4 implies that for sufficiently small $t \le t_0$ (with t_0 uniform in $n \ge 2$) we have

$$\left| \log \frac{w_t^{(n)}}{v_t^{(n)}} \right| \le \int_0^t \frac{\left| \Psi(v_t^{(n)}) - \Psi^{(\beta)}(v_t^{(n)}) \right|}{v_s^{(n)}} ds\,.$$

Using (8.9) and $v_s^{(n)} \le v_s \le C s^{-\frac{1}{\beta}}$ for small s (see (5.7)) we obtain

$$\left| \log \frac{w_t^{(n)}}{v_t^{(n)}} \right| \le C \left(\int_0^t (v_s)^{\beta-\alpha} ds + \int_0^t \log(v_s) ds \right) = O(t^{\frac{\alpha}{\beta}}) + O(t \log \frac{1}{t}). \tag{8.11}$$

Letting $n \to \infty$, we see that the same estimate holds also for $\left|\log \frac{w_t}{v_t}\right| = \left|\log \frac{v_t}{w_t}\right|$. In particular, $\lim_{t\to 0+} \log \frac{v_t}{w_t} = 0$, and so $\left|\frac{v_t}{w_t} - 1\right| \sim \left|\log \frac{w_t}{v_t}\right|$ for small t. We conclude that (8.2) holds since $\frac{\alpha}{\beta} > \frac{1}{1+\beta}$, completing the proof.

8.2 Limitations of robustness

In this section we provide an instructive counterexample, announced in both the Introduction and Remark 4.5. A careful reader will note that the just made arguments proving Theorem 4.4 are close to optimal, in that the power $\alpha = \frac{\beta}{1+\beta}$ should be critical for (8.2). Without making any general statements to this end, let us fix $\alpha \in (0, \frac{\beta}{1+\beta})$ and consider Λ such that

$$\Lambda(dy) = g(y)dy, \ y \in [0,1], \quad \text{where} \ \ g(y) := y^{-\beta}(1 + y^\alpha), y \in (0,1].$$

We keep the notation of the previous section, in particular v and w are as in (8.2). We will show that

$$t^{-\frac{1}{1+\beta}} \left(\frac{w_t}{v_t} - 1 \right) \text{ is unbounded as } t \to 0, \tag{8.12}$$

and that therefore the statement of Theorem 4.4 (b) cannot hold in this particular case.

By (8.4)–(8.8) (with $a = \frac{1}{2}$) we have

$$\Psi(q) - \Psi^{(\beta)}(q) = R_1(q) + O(q), \qquad q \geq 1.$$

Now R_1 can be written explicitly as

$$R_1(q) = q^2 \int_0^1 \int_0^r \int_0^{\frac{1}{2}} e^{-qyu} y^{\alpha-\beta} dy du dr.$$

Note that R_1 is again of the form (1.2) where Λ is given by $\Lambda_{\beta-\alpha}(dy) = y^{\alpha-\beta} \mathbb{1}_{[0,\frac{1}{2}]}(y) dy$. By (8.3), (8.6) and (8.8) with β replaced by $\beta - \alpha$ we obtain

$$\Psi(q) - \Psi^{(\beta)}(q) = \Psi^{(\beta-\alpha)}(q) + O(q) = Dq^{1+\beta-\alpha} + O(q), \qquad q \geq 1. \quad (8.13)$$

where D is a positive constant, that can be written explicitly.

Recall the expression for $\Psi^{(\beta)}$ given just after (8.2). It is easy to check that one can let $n \to \infty$ in (8.10), and obtain

$$\log \frac{w_t}{v_t} + C \int_0^t (w_s^\beta - v_s^\beta) ds = D \int_0^t v_s^{\beta-\alpha} ds + O(t) \qquad (8.14)$$

for all sufficiently small t, where C and D are positive constants (their exact value is not important for our purposes). As usual this is done via uniform (in small t and in n) control of the RHS in (8.10), see (8.11) for a similar argument. By (5.7) it follows that

$$\int_0^t v_s^{\beta-\alpha} ds \sim C_1 t^{\frac{\alpha}{\beta}}. \qquad (8.15)$$

Let us suppose that the function given in (8.12) is bounded near 0. Since $\alpha < \frac{\beta}{1+\beta}$, this implies that

$$\left| \frac{w_t}{v_t} - 1 \right| = o(t^{\frac{\alpha}{\beta}}), \quad \text{as} \quad t \to 0,$$

hence also

$$\left| \log \frac{w_t}{v_t} \right| \vee \left| \frac{v_t}{w_t} - 1 \right| = o(t^{\frac{\alpha}{\beta}}), \quad \text{as} \quad t \to 0. \qquad (8.16)$$

By an elementary application of Taylor's formula we have

$$\left| w_s^\beta - v_s^\beta \right| = \left| 1 - \left(\frac{v_s}{w_s} \right)^\beta \right| w_s^\beta \sim \beta \left| 1 - \frac{v_s}{w_s} \right| w_s^\beta \text{ as } s \to 0,$$

and since $w_s^\beta = K_1^\beta s^{-1}$, we conclude

$$\begin{aligned} \int_0^t \left| {w_s}^\beta - {v_s}^\beta \right| ds &\leq \beta K_1^\beta \int_0^t \frac{1}{s} \left| 1 - \frac{v_s}{w_s} \right| ds \\ &= \beta K^\beta \int_0^t o(s^{-1+\frac{\alpha}{\beta}})\, ds = o(t^{\frac{\alpha}{\beta}}). \end{aligned}$$

This together with (8.16) is in clear contradiction with (8.15) and (8.14). We conclude that the opposite of (8.2) must hold, or equivalently, that there must exist a positive constant c and a sequence of times $(t_n)_n$ such that $t_n \to 0$ and

$$\left|\frac{v_{t_n}}{w_{t_n}} - 1\right| \geq c(t_n)^{\frac{\alpha}{\beta}},$$

and joint with $\alpha \in (0, \frac{\beta}{1+\beta})$, this easily implies (8.12).

References

[1] D. J. Aldous (1999). Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists *Bernoulli*, 5(1):3–48.

[2] N. Berestycki (2009). *Recent progress in coalescent theory*. Ensaios Matematicos, Vol. 16.

[3] J. Berestycki, N. Berestycki, and V. Limic (2010). The Λ-coalescent speed of coming down from infinity. *Ann. Probab.*, 38(1):207-233.

[4] J. Berestycki, N. Berestycki and V. Limic (2012). Asymptotic sampling formulae for Lambda-coalescents. *AIHP*, to appear.

[5] J. Berestycki, N. Berestycki and V. Limic (2012). A small-time coupling between -coalescents and branching processes. *Ann. Appl. Probab.*, to appear.

[6] J. Bertoin (2006). *Random fragmentation and coagulation processes*. Cambridge University Press, Cambridge.

[7] J. Bertoin and J.-F. Le Gall (2006). Stochastic flows associated to coalescent processes III: Limit theorems. *Illinois J. Math*, 50, 147–181.

[8] P. Billingsley (1999). *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition.

[9] M. Birkner, J. Blath, M. Capaldo, A. Etheridge, M. Mhle, J. Schweinsberg, and A. Wakolbinger (2005). Alpha-stable branching and beta-coalescents. *Electron. J. Probab.*, 10:no. 9, 303–325.

[10] T. Bojdecki, L.G. Gorostiza, A. Talarczyk (2007). A long range dependence stable process and an infinite variance branching system. Ann. Probab. **35**, no. 2, 500527.

[11] J.-F. Delmas, J.-S. Dhersin, and A. Siri-Jegousse (2008) Asymptotic results on the length of coalescent trees. *Ann. Appl. Probab.*, 18, 997–1025.

[12] P. Donnelly, and T. Kurtz (1999). Particle Representations for Measure-Valued Population Models. *Ann. Probab.* **27**, 166–205.

[13] R. Durrett (2004). *Probability: theory and examples.* 3^{rd} edition. Duxbury advanced series.

[14] I. Iscoe (1986), A weighted occupation time for a class of measure-valued branching processes, *Probab. Th. Rel. Fields*, 71, 85-116.

[15] V. Limic (2010) On the speed of coming down from infinity for Xi-coalescent processes. *Electr. J. Probab.*, 15, 217–240.

[16] V. Limic (2011). *Processus de Coalescence et Marches Aléatoires Renforcées : Un guide à travers martingales et couplage.* Habilitation thesis (in French and English). Available at `http://www.latp.univ-mrs.fr/~vlada/habi.html`

[17] V. Limic (2012). Genealogies of regular exchangeable coalescents with applications to sampling. *Ann. IHP*, 48, No. 3, 706–720, 2012.

[18] S. Peszat and J. Zabczyk (2007). *Stochastic partial differential equations with Lvy noise*, volume 113 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge.

[19] J. Pitman (1999). Coalescents with multiple collisions. *Ann Probab.* **27**, 1870–1902.

[20] S. Sagitov (1999). The general coalescent with asynchronous mergers of ancestral lines. *J. Appl Prob.* **36** (4), 1116–1125.

[21] G. Samorodnitsky and M. S. Taqqu (1994). *Stable non-Gaussian random processes*. Stochastic Modeling. Chapman & Hall, New York.

[22] J. Schweinsberg (2000). A necessary and sufficient condition for the Λ-coalescent to come down from infinity. *Electron. Comm. Probab.* **5**, 1–11.

[23] J. Schweinsberg (2000). Coalescents with simultaneous multiple collisions. *Electron. J. Probab.* **5**, 1–50.

[24] J. Schweinsberg (2003). Coalescent processes obtained from supercritical Galton-Watson processes. Stoch. Proc. Appl. **106**, no. 1, 107139.